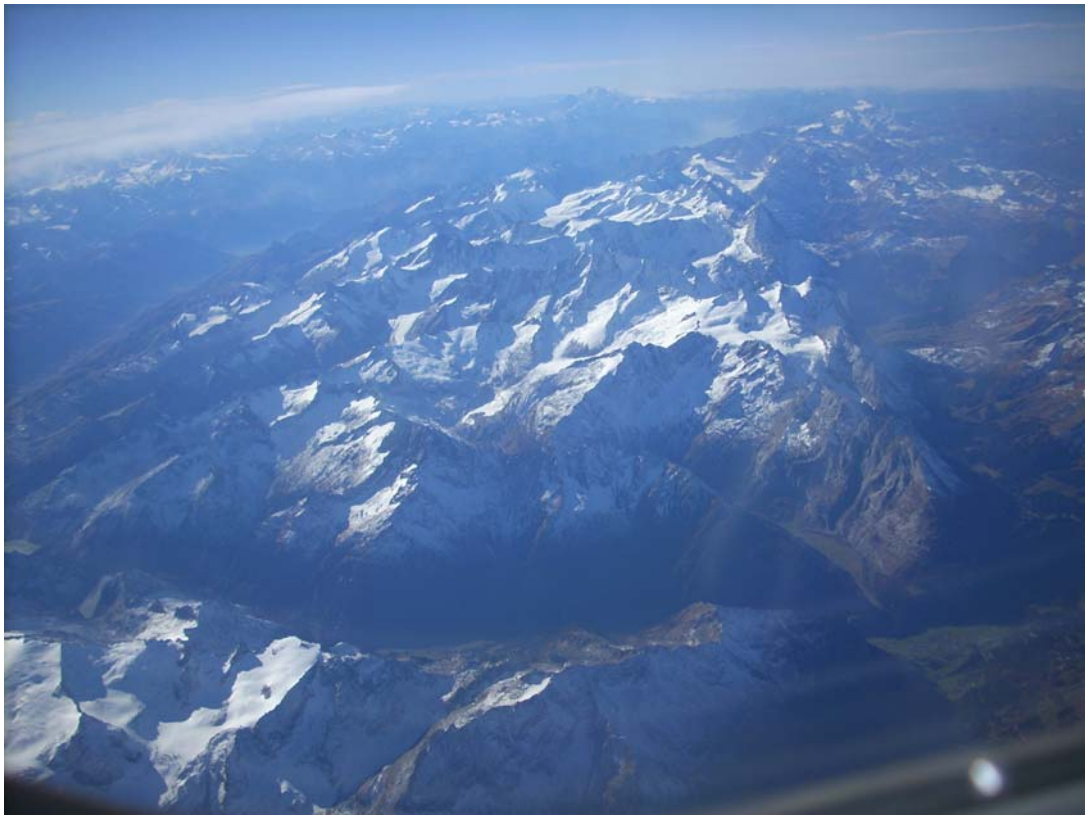
JRC Scientific and Technical Reports

Demonstration of Equivalence of Ambient Air Analytical Method

PROCEEDINGS of the WORKSHOP on Demonstration of Equivalence
2-4 May 2007 Ispra (VA) Italy

GUIDELINE on Demonstration of Equivalence
Novembre 2005



Edited by P. Pérez Ballesta

EUR 23216 EN - 2008

JRC EUROPEAN COMMISSION

ies Institute for Environment and Sustainability

The mission of the Institute for Environment and Sustainability is to provide scientific-technical support to the European Union's Policies for the protection and sustainable development of the European and global environment.

European Commission
Joint Research Centre
Institute for Environment and Sustainability

Contact information
Address: TP 441 JRC 21027-Ispra (VA) Italy
E-mail: pascual.ballesta@jrc.it
Tel.: +390332785322
Fax: +390332789364

http://ies.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

Legal Notice
Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number (*):

00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server http://europa.eu/

JRC 42520

EUR 23216 EN
ISBN 978-92-79-08205-4
ISSN 1018-5593
DOI 10.2788/63342

Luxembourg: Office for Official Publications of the European Communities

© European Communities, 2008

Reproduction is authorised provided the source is acknowledged

Printed in Italy

TABLE OF CONTENTS

FOREWORD **5**

AGENDA OF THE WORSHOP **6**

SUMMARY AND OPEN DISCUSSION OF THE WORKSHOP ON DEMONSTRATION OF EQUIVALENCE **8**

ABSTRACT BOOK **12**

- Introduction to Equivalence. P. Pérez Ballesta, (DG-JRC) 13
- Future legislation and the role of the Institutions in the definition of equivalence
- Kobe (DG-Env) 14
- The importance of the standarisation for the implementation of the EU Directive. K. Saunders (Keris Ltd., UK) 15
- The role of the DG JRC in the harmonization of air quality measurements in Europe. A. Borowiak (DG-JRC) 17
- Demonstration of equivalence: Procedure, test and evaluation. T. Hafkenscheid (NMi, NL) 18
- Comparison of different regression methods in equivalence demonstration. R. Beier (LANUV, NRW, D) 19
- New issues related to the PM10 and $PM_{2.5}$ standards. A. van der Meulen (RIVM, NL) 20
- Experiences with equivalence testing in Germany. U. Pfeffer (LANUV, NRW, D) 21
- Equivalence of PM10 measurements in the Netherlands national air quality monitoring network. R. Hoogerbrugge, R. Beijk (RIVM, NL) 23
- Practical testing of equivalence demonstration for PM automated monitoring methods: The Experience in France. O. Le Bihan (LCSQA-INERIS, F) , F. Mathé (LCSQA, Mines de Douai, F), H. Marfaing (Airparif, F), D. Robin (Airmaraix, F) 24
- Example of evaluation of equivalence for the NO2 membrane-closed Palmes tube and O3 radial diffusive sampler. M. Gerboles and D. Buzica (DG-JRC), H. Plaisance (LCSQA, EMD, F) 27
- Particulate matter equivalence in the United Kingdom. J. Dixon (DEFRA, UK) 28
- PM10 equivalence tests in Switzerland: relationship between a TEOM-FDMS and the reference method at a suburban site. Ch. Hueglin, R. Gehrig (EMPA, CH) 30
- Field campaigns to determine the correction factors in the Spanish atmospheric networks. S. García Dos Santos, P. Díez Hernández , R. Fernández Patier (ISCIII, E) 31
- Demonstration of equivalence for PM monitors in Flanders (Belgium). J. Vercauteren, C. Matheeussen, E. Roekens (FEA, BE) 34
- Qualifying a $PM_{10}/PM_{2.5}$ method as transfer standard or for automated standard: Experience with the swam 5-a. L. Bertrand, G.Gérard and S.Fays (ISSeP, BE) 36
- Atmospheric networks. Equivalence of PM10 in Slovenia. T.Bolte, P.Pavli (EARS, SLO) 37
- Non seasonable variability and the effect of logging averages on TEOM and FDMS measurements made during the United Kingdom PM_{10} and $PM_{2.5}$ equivalence trials. D. Harrison, J. Booker, R. Maggs (BV HS &E, UK) 38
- Experience on demonstration of equivalence on $PM_{10}/PM_{2..5}$ 3 issues. D. de Jonge (GGD, NL) 39

PAPERS **41**

- COMPARISON OF REGRESSION METHODS IN EQUIVALENCEDEMONSTRATION. Reinhold Beier 42
- NEW ISSUES RELATED TO THE PM10 AND PM2.5 STANDARDS: TOWARDS A ROBUST CONTROL OF UNCERTAINTY. A. van der Meulen. 49
- EXPERIENCES WITH EQUIVALENCE TESTING IN GERMANY. Ulrich Pfeffer, Thorsten Zang, Reinhold Beier and Eva-Maria Doutnik. 54
- EQUIVALENCE OF PM10-MEASUREMENTS IN THE DUTCH NATIONAL AIR QUALITY MONITORING NETWORK Ronald Hoogerbrugge, Ruben Beijk. 61
- PRACTICAL TESTING OF EQUIVALENCE DEMONSTRATION FOR PM AUTOMATED MONITORING METHODS: THE EXPERIENCE IN FRANCE. O. Le Bihan, F. Mathé, H. Marfaing, D. Robin 67
- EQUIVALENCE OF DIFFUSIVE SAMPLERS TO REFERENCE METHODS: THE OZONE RADIAL DIFFUSIVE SAMPLER AND MEMBRANE-CLOSED PALMES TUBE EXAMPLES. Michel Gerboles, Daniela Buzica, Hervé Plaisance, A. Piechocki, G. Galloo 72
- PM10 EQUIVALENCE TESTS IN SWITZERLAND: RELATIONSHIP BETWEEN A TEOM-FDMS AND THE REFERENCE METHOD AT A SUBURBAN SITE. Ch. Hueglin, R. Gehrig, Beat Schwarzenbach. 82
- NON SEASONAL VARIABILITY AND THE EFFECT OF LOGGING AVERAGES ON TEOM AND FDMS MEASURMENTS MADE DURING THE UNITED KINGDOM PM_{10} and $PM_{2.5}$ EQUIVALENCE TRIALS. David Harrison, Jeff Booker and Richard Maggs. 87

ANNEX
DEMONSTRATION OF EQUIVALENCE OF AMBIENT AIR MONITORING METHODS (November 2005) 92

FOREWORD

The EU directives on ambient air quality define a series of reference methods, which have to be used by the Member States to monitor ambient air quality. These methods have been standarised by CEN and their analytical uncertainties have been determined to be in line with the data quality objectives of the directives. The use of equivalence methods for reporting ambient air monitoring data to the Commission should be guaranteed by a robust procedure, which is capable of evaluating the uncertainty associated with the candidate method under similar operational field conditions.

The grade of complexity of the procedure for demonstration of equivalence depends on how much the candidate method differs from the reference method. Such a variation ranges from a modification of the minor part of the standardised method to a method that could work with a completely different principle. The "Guideline on Demonstration of Equivalence" deals with these matters, first by defining equivalence as "*…a method meeting the Data Quality Objectives for continuous or fixed measurements specified in the relevant air quality directive*" and, secondly, by establishing a procedure for demonstration of equivalence based on specific programmes for laboratory and field tests with the ultimate goal of determining the uncertainty of the candidate method in comparison to the reference method.

This publication collates the proceedings of the Workshop on Demonstration of Equivalence between Ambient Air Monitoring Methods, which was held at the JRC in Ispra (VA) from 2 - 4 May 2007 and the major discussions vis-à-vis the last version of the "Guideline on Demonstration of Equivalence" (included as an annex).

During the workshop in Ispra, a series of relevant items were discussed including the same approaching principle, practical discussions on experimental design, QAQC requirements and other statistical approaches. All this information has been collected in the current proceedings and the different presentations slides can be found on the following website of the Joint Research Centre: http://ies.jrc.cec.eu.int/626.html.

It should be understood that the articles included in these proceedings do not represent any official position of the European Commission, rather the point of view of the corresponding authors. Theses proceedings provide an extensive discussion on the current limitations of the guideline on Demonstration of Equivalence and, consequently, give some space for possible improvements. The Guideline is reinforced through extensive application and use by the reference laboratories from Member States.

Ispra, 7 January 2008

Pascual Pérez Ballesta

AGENDA

WORKSHOP ON EQUIVALENCE OF ANALYTICAL METHODS

JRC, Ispra (VA) Italy, 2-4 May 2007

The slides from the workshop can be downloaded at the website: http://ies.jrc.cec.eu.int/626.html

CURRENT STATE OF THE ART

Chair: Peter Woods (NPL, UK), Rapporteurs: U. Pfeffer (LANUV, NRW, D)
2 May 2007. Wednesday afternoon , 14:00 — 17:00

Welcome and introduction to Equivalence
P. Pérez Ballesta, (DG-JRC)
Future legislation and the role of the Institutions in the definition of equivalence
A. Kobe (DG-Env)
The importance of the standarisation for the implementation of the EU Directive
K. Saunders (Keris Ltd., UK)
The role of the DG JRC in the harmonization of air quality measurements in Europe
A. Borowiak (DG-JRC)
Implementation of European Standards and Guidelines: Certification & Type-Approval versus Equivalence & Method Validation
P. Woods (NPL, UK)

FOCUSING ON THE ISSUE
Chair: P. Quincey (NPL, UK) , Rapporteurs: A. Borowiak (DG-JRC)
2 May 2007. Wednesday afternoon , 14:00 — 17:00

Demonstration of equivalence: Procedure, test and evaluation
T. Hafkenscheid (NMi, NL)
3 May 2007. Thursday morning, 9:00 — 12:30
Comparison of different regression methods in equivalence demonstration
R. Beier (LANUV, NRW, D)
New issues related to the PM10 and $PM_{2.5}$ standards
T. van der Meulen (RIVM, NL)
Demonstration of equivalence: Comments and experiences
T. Hafkenscheid (NMi, NL)

EXPERIENCES IN DIFFERENT MEMBER STATES
Chair: T. Hafkenscheid (NMi, NL), Rapporteurs: M. Froehlich (Umweltbundesamt GmbH, A)
3 May 2007. Thursday morning, 9:00 — 12:30

Experiences with equivalence testing in Germany
U. Pfeffer (LANUV, NRW, D)
Equivalence of PM10 measurements in the Netherlands national air quality monitoring network.
R. Hoogerbrugge, R. Beijk (RIVM, NL)

Practical testing of equivalence demonstration for PM automated monitoring methods: The Experience in France.
O. Le Bihan (LCSQA-INERIS, F) , F. Mathé (LCSQA, Mines de Douai, F), H. Marfaing (Airparif, F), D. Robin (Airmaraix, F)

3 May 2007. Thursday afternoon, 14:00—17:00

Example of evaluation of equivalence for the NO2 membrane-closed Palmes tube and O3 radial diffusive sampler.
M. Gerboles and D. Buzica (DG-JRC), H. Plaisance (LCSQA, Mines de Douai, F)
Particulate matter equivalence in the United Kingdom.
J. Dixon (DEFRA, UK)
PM10 equivalence tests in Switzerland: relationship between a TEOM-FDMS and the reference method at a suburban site
Ch. Hueglin, R. Gehrig (EMPA, CH)

POSTER DISCUSSION

Field campaigns to determine the correction factors in the Spanish atmospheric networks
S. García Dos Santos, P. Díez Hernández , R. Fernández Patier (ISCIII, E)
Demonstration of equivalence for PM monitors in Flanders (Belgium)
J. Vercauteren, C. Matheeussen, E. Roekens (FEA, BE)
Qualifying a $PM_{10}/PM_{2.5}$ method as transfer standard or for automated standard: Experience with the swam 5-a. L. Bertrand, G.Gérard and S.Fays (ISSeP, BE)
Atmospheric networks. Equivalence of PM10 in Slovenia
T.Bolte, P.Pavli (EARS, SLO)
Non seasonable variability and the effect of logging averages on TEOM and FDMS measurements made during the United Kingdom PM_{10} and $PM_{2.5}$ equivalence trials
D. Harrison, J. Booker, R. Maggs (BV HS &E, UK)
Experience on demonstration of equivalence on $PM_{10}/PM_{2..5}$ 3 issues
D. de Jonge (GGD, NL)

OPEN DISCUSSION

Chair: T. Hafkenscheid (NMi, NL)

4 May 2007. Friday morning 9:00 — 12:00
(Open discussion) :
Summary of previous sessions: rapporteurs
Future challenges: new pollutants and techniques: What is missing?
Implementation of the equivalence report
Inputs, corrections and final publication of the equivalence report
Current status of the equivalence report

SUMMARY AND OPEN DISCUSSION OF THE WORKSHOP ON DEMONSTRATION OF EQUIVALENCE

Theo Hafkensheid. NMi van Swinden laboratory. NL – Delft. thafkenscheid@nmi.nl

Objective

The main objective of the workshop was to generate information for the improvement/revision of the Guidance Report for the Demonstration of Equivalence, aiming where possible at simplification and streamlining of procedures, and at a broad acceptance of conclusions.

Overview of discussion items

Broadly, the items discussed during the workshop can be categorized as follows.

- The general approach to equivalence demonstration, including interpretation of elements of the guidance report
- The practical approach to equivalence demonstration, with a main focus on particulate matter
- (Statistical) evaluation approaches used
- Ongoing QA/QC
- Practical experiences in different member states, including the use of new approaches
- Dissemination/sharing of information.

Summary of discussion and conclusions

Main focus

The main focus of the workshop has been on (items related to) the equivalence of automated continuous methods for measuring particulate matter. As a part of the discussion also items not strictly related to equivalence but relevant for its establishmenthave been discussed:

- The intrinsic uncertainty of the current realizations of the reference method for measuring PM (e.g. selection of type and brand of filter material; weighing conditions)
- Type approval of automated continuous methods.

Further discussion of these items is within the scope of the current activities of CEN/TC264 WG 15.

General approach to equivalence demonstration

Establishment of equivalence

The national competent authority – based on a validation study, considering the scope of the application of the equivalent method (e.g. site conditions, concentration levels, environmental and meteorologocal conditions) – authorizes the use of the equivalent method at the national level. The declaration of equivalence will include its scope (see above), and the period of validity of the claim.

Generalization/mutual acceptance/recognition of equivalence claims

The potential for a mutual acceptance of equivalence demonstrations between member states is limited by the scopes of the individual claims. For methods for which equivalence has been demonstrated over a broad range of conditions/scopes (see above) a broader claim of equivalence may be considered. The

AQUILA forum may have an important role in this by providing advice on the generalization of claims.

Definition of a candidate equivalent method, particularly for PM

A candidate method may include corrections other than those derived from the evaluation of equivalence, e.g. corrections for effects of temperature and relative humidity if the uncertainties resulting from the corrections are included in the uncertainty budget of the candidate method.

In addition, more that one correction may be used for one candidate method if the scope of the associated equivalence claims is well defined, and if separate ongoing QA/QC suitably includes all differentiations.

A new approach is currently tested in the UK and France. The approach comprises a "dynamic" "remote" correction of hourly TEOM data using correlations with a regionally centered FDMS-TEOM. In this approach the candidate method is a "model system" rather than a method.
This approach should be accompanied with appropriate estimations of the uncertainties associated with these corrections.

In any case, the current equivalence protocol does not cover such approaches, and needs updating to include these new developments.

Practical approach to equivalence demonstration

Use of an equivalent method instead of a realization of the reference method

It has been proposed to use an already equivalent method as "reference" for comparison with a new candidate method. Here a distinction shall be made between a variation on the realization of the reference method (e.g. a sampler with automatic sample changer) and a different methodology.

In the former case the use of the equivalent method is acceptable as long as the effect of the variation is known and not significant.

In the latter case the concept of comparison with a "reference" may be invalidated because the measurement results of the different methodology will almost certainly have a bias towards the results of the reference method (it is assumed that the reference method is unbiased by definition). This requires a careful reconsideration of the statistics to be applied for the uncertanty evaluation. Moreover, the uncertainty of the equivalent method should be fully propagated into that of the new candidate method, most probably resulting in high uncertainties.

Obligation to evaluate equivalence for various data subsets (PM)

Questions have been raised as to the necessity to

- separately evaluate uncertainty for PM candidate methods for 7 different datasets, including subsets based on concentration ranges
- to pass the evaluation criteria for all separate datasets.

Sometimes practical limitations (e.g. low concentrations for specific sites) will prevent the full evaluation required.
This issue requires a reconsideration of the requirements of the current evaluation protocol.

Application of corrections

Questions have been raised as to whether corrections of slope and/or intercept should be applied when

- the slope / intercept are significantly different from 1 / 0 at the 95% confidence level; it has been proposed rather to use fixed ranges for each (cf. EN 12341)
- the uncertainty of the candidate method already fulfills the equivalence criterion.

In principle, ISO-GUM states that corrections shall always be applied in case of significant differences. In practice this may only be useful when the resulting uncertainty is positively affected.

Another factor to consider may be the uncertainty of the reference method itself.

For PM an additional issue may be the effect on the number of exceedances of daily limit value, although the uncertainty associated with the determination of this number is high already.

Statistical procedures

Removal of outliers
It has been proposed not to base the removal of outliers solely on technical grounds, but rather to use a statistical level of significance (e.g. 99%). An argument for using a statistical approach may be that in a large dataset a certain fraction of data from the manual standard method will be "aberrant" beyond demonstrability anyway.

Use of orthogonal regression forced through (0,0)
It has been proposed to use orthogonal regression forced through (0,0) when the intercept obtained by the general orthogonal regression procedure is not significantly different from zero.
In order to be able to accept this approach the effects on the slope of the regression (the calibration factor) should be understood. Possibly limitations should be set to the magnitude of the intercept.

A further question is how the uncertainty of the corrected results should be calculated.
At minimum the uncertainty of the correction factor shall be included in the uncertainty budget.

Use of different regression techniques

The introduction of a different regression method based on a criterion of symmetry along the correlation line instead of minimizing the sum of residual values has been proposed. Although it has been shown that when applying this criterion the Exploratory Regression method may have a better performance than the orthogonal regression technique, it is unclear whether this criterion provides a better correlation line for equivalence purposes because of an increased number of outliers and a potential underestimation of the uncertainty of the candidate method.

Ongoing QA/QC

The practical application of an equivalent method should always be accompanied by some form of ongoing QA/QC.
In the case of PM this may consist of periods of parallel measurements performed using a realization of the reference method.

It has been suggested to link the intensity of the QA/QC to the uncertainty level of the equivalent method.
If one candidate method is applied with different corrections the ongoing QA/QC should be such that all differentiations within the candidate method are suitably covered.

Experiences in Member States

A number of Member States presented their experiences – both orally and through posters – with the application of the approach to equivalence demonstration laid down in the current Guidance Report.

Dissemination of information

During the workshop it has been concluded that currently a "threshold" exists for sharing and exchanging information on equivalence trials and experiences. It has been suggested to open a website specifically for the exchange of information on equivalence.
Apart from serving as a medium for sharing information and discussion, the site may contain a list of equivalent methods plus scopes that have been "approved" by the AQUILA forum.

ABSTRACT BOOK

- Introduction to Equivalence. P. Pérez Ballesta, (DG-JRC)
- Future legislation and the role of the Institutions in the definition of equivalence
- Kobe (DG-Env)
- The importance of the standarisation for the implementation of the EU Directive. K. Saunders (Keris Ltd., UK)
- The role of the DG JRC in the harmonization of air quality measurements in Europe. A. Borowiak (DG-JRC)
- Demonstration of equivalence: Procedure, test and evaluation. T. Hafkenscheid (NMi, NL)
- Comparison of different regression methods in equivalence demonstration. R. Beier (LANUV, NRW, D)
- New issues related to the PM10 and $PM_{2.5}$ standards. T. van der Meulen (RIVM, NL)
- Demonstration of equivalence of Ambient Air Monitoring Methods. T. Hafkenscheid (NMi, NL)
- Experiences with equivalence testing in Germany. U. Pfeffer (LANUV, NRW, D)
- Equivalence of PM10 measurements in the Netherlands national air quality monitoring network. R. Hoogerbrugge, R. Beijk (RIVM, NL)
- Practical testing of equivalence demonstration for PM automated monitoring methods: The Experience in France. O. Le Bihan (LCSQA-INERIS, F) , F. Mathé (LCSQA, Mines de Douai, F), H. Marfaing (Airparif, F), D. Robin (Airmaraix, F)
- Example of evaluation of equivalence for the NO2 membrane-closed Palmes tube and O3 radial diffusive sampler. M. Gerboles and D. Buzica (DG-JRC), H. Plaisance (LCSQA, Mines de Douai, F)
- Particulate matter equivalence in the United Kingdom. J. Dixon (DEFRA, UK)
- PM10 equivalence tests in Switzerland: relationship between a TEOM-FDMS and the reference method at a suburban site. Ch. Hueglin, R. Gehrig (EMPA, CH)
- Field campaigns to determine the correction factors in the Spanish atmospheric networks
- S. García Dos Santos, P. Díez Hernández , R. Fernández Patier (ISCIII, E)
- Demonstration of equivalence for PM monitors in Flanders (Belgium). J. Vercauteren, C. Matheeussen, E. Roekens (FEA, BE)
- Qualifying a $PM_{10}/PM_{2.5}$ method as transfer standard or for automated standard: Experience with the swam 5-a. L. Bertrand, G.Gérard and S.Fays (ISSeP, BE)
- Atmospheric networks. Equivalence of PM10 in Slovenia. T.Bolte, P.Pavli (EARS, SLO)
- Non seasonable variability and the effect of logging averages on TEOM and FDMS measurements made during the United Kingdom PM_{10} and $PM_{2.5}$ equivalence trials
- D. Harrison, J. Booker, R. Maggs (BV HS &E, UK)
- Experience on demonstration of equivalence on $PM_{10}/PM_{2..5}$ 3 issues. D. de Jonge (GGD, NL)

INTRODUCTION TO EQUIVALENCE

P. Pérez Ballesta
European Commission. Joint Research Centre. Institute. Ispra (VA) Italy.
pascual.ballesta@jrc.it

The harmonization of the air quality measurements in Europe attempts to guarantee the comparability of air quality data coming from different places, carried out by different laboratories, instrumentation, methods and in different environmental conditions.
The European Commission through the Air Quality Directives is trying to formalize this harmonization process by defining reference methods, sampling criteria and data quality objectives. In this context the concept of equivalence is very much linked to the objective of guaranteeing the same quality of measurements of the reference method.
Analytical quantification is based on the measure of a physical-chemical property, which is directly or indirectly related to concentration. This covers direct continuous measurements and any analytical method that may include concentration and separation of its compounds prior to the measurement.
The need of an equivalence protocol is inherent to the process of quantification as different methods could measure different properties and be related in different ways to concentration. Furthermore, environmental variables may also affect any step of the quantification process.
What is understood as “equivalent”?
By assuming that an equivalence method is a method that meets the data quality objectives of the directive, in other words, that meets the same uncertainty criteria of the reference method, it is not sufficient that the equivalence method provides averaged results comparable to the reference method, it has to be precise enough to guarantee a limited uncertainty of the results.
That implies that the test of equivalence cannot only be based on a null hypothesis of non statistical differences with respect to a reference method. But the equivalence should be established after calculation of the uncertainty associated with the candidate method.
Method uncertainty can be calculated according to ISO GUM. Such an uncertainty calculation shall cover the range of environmental conditions in which the method is applied.
The whole process may include laboratory tests under repeatability and reproducibility conditions as well as other tests under field conditions. Furthermore, for some methods where a primary standard is not available candidate methods can only be evaluated by comparison to a reference method under field conditions.
A statistical test has been proposed in order to evaluate the uncertainty of a candidate method when compared to a reference one. The uncertainty derived from this test for the candidate method should be independent of the uncertainty of the reference method and comparable to the uncertainty calculated by pooling all possible contributions to the uncertainty budget.

FUTURE LEGISLATION AND THE ROLE OF THE INSTITUTIONS IN THE DEFINITION OF EQUIVALENCE

A. Kobe
European Commission, DG Environment. Clean Air and Transport Unit.
andrej.kobe@ec.europa.eu

Air pollution has very strong adverse health effects. According to the latest scientific and health evidence, presented in the Commission Communication on Thematic Strategy on Air Pollution COM(2005) 446, only exposure to fine particulate matter $PM_{2.5}$ in ambient air is responsible for the reduction of the statistical life expectancy of average EU citizen by more than 8 months. In its proposal the Commission thus introduces specific environmental standards for fine particulate matter $PM_{2.5}$ in ambient air. Their implementation should significantly contribute to reaching objective of the Thematic Strategy on Air Pollution which is to reduce the number of life years lost in Europe due to exposure to particulate matter by 47% in the period between 2000 and 2020.

Following on from the Commission initiative on "better regulation", the Commission proposal for a directive on ambient air quality and cleaner air for Europe further merges the provisions of the framework and the three daughter directives on ambient air quality together with the Council decision on the Exchange of Information into a single directive with the intention of simplifying, streamlining and reducing the volume of existing legislation. In addition the proposal revises the existing provisions so as to incorporate the experience of the Member States. For that purpose the proposal a) introduces specific monitoring requirements and new environmental objectives for fine particulate matter $PM_{2.5}$, b) provides some flexibility in the implementation by allowing, under specific conditions to be approved by the Commission, prolongation of the attainment dates for certain limit values such as for the particulate matter PM10 and nitrogen dioxide, and c) enables the Member States to focus their efforts by allowing deduction of natural contributions when assessing compliance with the limit values.

The Commission proposal is currently in the co-decision by the European Parliament and the Council. Both institutions have proposed changes to the commission proposal. The differences will be further addressed in the so-called second reading which is expected to start in September 2007.

The Commission supports the Council political agreement, which has proposed a number changes, but had maintained the important balance between the increased flexibility of implementation and the protection of public health. Assessment of some of the Parliaments amendments is however that they would weaken the ambition level of the original Commission proposal and might introduce delay in the implementation of measures. Discussion continues as all Institutions share the objectives as described in the Thematic Strategy on Air Pollution. If the agreement is reached in the second reading the Directive may enter into force in the beginning of 2008.

The Commission proposal as well as the new Council text both maintain the concept of quality assurance and quality control as well as the insurance of comparability of measurements through the equivalence requirement, set by the existing Directives. The language has been streamlined to make more explicit reference to traceability, QA/QC of data collection and reporting, and the requirement for the accreditation of the national reference laboratories. These are given also explicit coordination task of the appropriate realization of the reference methods as well as the demonstration of equivalence at the national level, coordination of Community QA/QC programmes at the national level as well as the participation in Community-wide intercomparisons.

According to the new provisions, the demonstration of equivalence report may be required by the Commission in order to be able to assess the compliance with the provisions regarding use of the non-reference methods for the ambient air quality assessment. When assessing the acceptability of the report, the Commission will make reference to its guidance on the demonstration of equivalence. This guidance will also include the ways to be followed when approximating equivalence through use of interim factors.

THE IMPORTANCE OF STANDARDISATION FOR THE IMPLEMENTATION OF EU DIRECTIVES

K. J. Saunders
KERIS LTD Heckfield, Hampshire, UK. SaundersKJ@KerisLtd.Com.

A number of Ambient Air Directives have been implemented in last 10 years. The framework directive, Council Directive 96/62/EC "ambient air quality assessment and management". The first daughter directive, Council Directive 1999/30/EC, "limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air". The second daughter directive, Council Directive 2000/69/EC, "limit values for benzene and carbon monoxide in ambient air". The third daughter directive, Council Directive 2002/3/EC, "ozone in ambient air" The fourth daughter directive, Council Directive 20004/107/EC, "arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air". A further directive has been proposed. Directive Proposal 2007/xxx/EC, "ambient air quality and cleaner air for Europe". The proposal aims to revise and combine in one directive the framework directive and the first, second and third daughter directives. It is intended that the fourth daughter directive will remain in force for the time being.

CEN TC 264 has been mandated to write standards, which fulfil the requirements in these directives and some of these standards have been designated as reference methods by European Commission.

The first standard produced was for PM10 from Working Group 7:

EN12341:1999 Air Quality – Determination of the PM10 fraction of suspended particulate matter – Reference method and field procedure to demonstrate reference equivalence of measurement methods.

This was followed with standards from Working Group 12:

EN14211:2005 standard "for the measurement of the concentrations of nitrogen dioxide and nitrogen monoxide by chemiluminescence".
EN14212:2005 standard "for the measurement of the concentrations of sulphur dioxide by ultraviolet fluorescence".
EN 14625:2005 standard "for the measurement of the concentrations of ozone by ultraviolet photometry".
EN 14626:2005 standard "for the measurement of the concentrations of carbon".

Working Group 13 produced standard methods for the measurement of benzene concentrations.:

EN14662-1 Part 1: "pumped sampling followed by thermal desorption and gas chromatography".
EN14662-2 Part 2: "pumped sampling followed by solvent desorption and gas chromatography".
EN14662-3 Part 3: "automated pumped sampling with in situ gas chromatography".
EN14662-4 Part 4: "diffusive sampling followed by thermal desorption and gas chromatography".
EN14662-5 Part 5: "diffusive sampling followed by solvent desorption and gas chromatography".

Working group 14 produced a standard for heavy metals:

EN 14902:2005 standard "for the measurement of Pb, Cd, As and Ni in the PM10 fraction of suspended particulate matter".

Working Group 15 produced a standard for $PM_{2.5}$:

EN14907 standard “gravimetric measurement method for the determination of the PM 2.5 fraction of suspended particulate matter”.

Working Group 21 is working on a standard for BaP:

prEN 15549 standard “for the measurement of the concentration of benzo[a]pyrene”. This standard has been for CEN enquiry.

Other standards being produced include, the deposition of heavy metals, the deposition of PAH’s, a revised standard for PM10 and a standard for gaseous mercury.

THE ROLE OF DG-JRC IN THE HARMONISATION OF AIR QUALITY MEASUREMENTS IN EUROPE

A. Borowiak, M. Gerboles, F. Lagler, L. Marelli. and E. De Saeger
European Commission – DG Joint Research Centre. Institute for Environment and Sustainability. Transport and Air Quality Unit. T.P. 441, I – 21020 Ispra (VA)

The need for comparable data from air quality measurements has already been confirmed in the 1996's Air Quality Framework Directive [1]. The Directive is asking for an air quality assessment through common methods and criteria. Those criteria regarding the assessment are given and in the following Daughter Directives on the different air pollutants described. The Proposal for a Directive of the European Parliament and of the Council on ambient air quality and cleaner air for Europe [2] is going a step further and into more detail: Responsibility has to be taken over by the Member States for type approval of measurement systems, accuracy of measurements, traceability of measurements, a properly implemented QA/QC programme in the monitoring network, analysis of assessment methods and the co-ordination of the European Commission's quality assurance programmes. National Air Quality Reference Laboratories shall be accredited according to ISO 17025 at least for the reference measurement methods. The European Commission's Joint Research Centre in Ispra is supporting the Air Quality Directive's efforts since many years via various actions. Examples will be given for the activities on the harmonisation of measurement strategies, contributions in the field of standardization of air quality assessment, and quality assurance programmes at European scale. The latter will be described in detail through activities in the field of ozone and other inorganic gaseous pollutants, volatile organic compounds and ozone precursors, and the currently on-going European programme for particulate matter, including heavy metals in PM.

[1] Council Directive 96/62/EC of 27 September 1996 on ambient air quality assessment and management

[2] Proposal for a Directive of the European Parliament and of the Council on ambient air quality and cleaner air for Europe, COM (2005) 447 (21.09.2005)

DEMONSTRATING EQUIVALENCE OF AMBIENT AIR MONITORING METHODS

Theo Hafkenscheid
NMi van Swinden laboratory. NL – Delft. thafkenscheid@nmi.nl

European Ambient Air Quality Directives (1999/30/EC, 2000/69/EC, 2002/3/EC, 2004/147/EC) specify reference methodologies for the measurement of various air pollutants. These methodologies have been standardized by Working Groups of CEN TC 264 'Air Quality'. However, their use is not mandatory. The Directives allow Member States to use alternative methods provided that equivalence of results with those obtained using the CEN Standard Methods is demonstrated.
In order to provide guidance to those who want to use alternative methods to the CEN Standard Methods, a Working Group of the European Commission has drafted a guidance report for the performance of equivalence testing and subsequent evaluation of test results. In this presentation the principles and procedures for the demonstration of equivalence of ambient air monitoring methods will be outlined, and some examples will be given for equivalence testing, with emphasis on automated methods for the monitoring of particulate matter.

COMPARISON OF REGRESSION METHODS IN EQUIVALENCE DEMONSTRATION

R. Beier
Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, D-45133 Essen
reinhold.beier@lanuv.de, Tel +49-201-79951474, Fax +49-201-7995-1575

Regression lines obtained in evaluating inter-comparison exercises of PM10-measuring methods are validated by comparison with alternative regression lines. Validation is based on a simple and robust approach to objective regression diagnostics. Figures of merit known from Ordinary Least Squares Regression (OLS) like residual standard deviation and the coefficient of determination turn out inappropriate within this context of errors in both variables. Instead, three robust criteria are used to assess basic statistical properties such as randomness and symmetry of the distribution of residuals as well as coverage of the 95%-uncertainty range. These performance criteria are used to evaluate the fitness for purpose of regression lines. This procedure was used to compare regression lines obtained in evaluation of inter-comparison exercises of automatic PM10-monitors with a collocated High-Volume Filter-Sampler by means of Ordinary Least Squares Regression, Orthogonal Regression and Explorative Regression [1]. In this exercise, OLS-Regression and Explorative Regression were found advantageous in about 30 percent of the evaluated cases. Orthogonal Regression turned out preferable in the 40 percent of the evaluated data sets.

[1] Beier, R.: Evaluation of method intercomparison tests by Explorative Regression, Gefahrstoffe – Reinhaltung der Luft, 67(2007), 49-54.

NEW ISSUES RELATED TO THE PM10 AND $PM_{2.5}$ STANDARDS: TOWARDS A ROBUST CONTROL OF UNCERTAINTY

A. van der Meulen
Convenor CEN/TC264/WG15: $PM_{2.5}$ National Institute for Public Health and the Environment RIVM Bilthoven / Netherlands

The current focus on equivalence almost "automatically" calls for:
a closer look into the performance of the EU Standard Methods for PM (EN12341 for PM10 and EN14907 for $PM_{2.5}$), and consideration of the possibility to introduce standards for Automatic Monitoring Systems AMS.

The issue of equivalence, i.e. whether a non-standard method gives results equivalent for regulatory purposes to EU Standard Methods, has been addressed by a separate EC working group for all pollutants in a comprehensive and uniform way.
For the evaluation of the "(lack of) comparability" between a non-standard and the standard measurement method (expressed as uncertainty) it is assumed that the relationship between the measurement results from both methods can be described by a linear relation, to be established by orthogonal regression. Subsequently, from this orthogonal regression the uncertainty from the non-standard method can be inferred.
As a first remark, obviously the "lack of comparability" between a non-standard and the Standard Method (SM) is determined inter alia by the performance of the SM. A lower SM uncertainty does improve the reliability of the equivalence claim.
Secondly, for most Automatic Monitoring Systems AMS a good comparability is observed at two different specific sites and ambient conditions (R values of 0,94 or higher).

Prompted by this observation, it could be worthwhile to consider standardization of AMS, including measurement ànd data treatment. For, it should be noted that for the greater part EU monitoring networks are employing Automated Monitoring Systems.

EXPERIENCES WITH EQUIVALENCE TESTING IN GERMANY

Ulrich Pfeffer
Landesamt für Natur, Umwelt und Verbraucherschutz (LANUV) Nordrhein-Westfalen, Essen (Germany) ulrich.pfeffer@lanuv.nrw.de, Tel. +49-201-7995 1264, Fax +49-201-7995 1575

Experience across Europe shows that most 'classical' continuous PM10 monitoring methods underestimate real concentrations by approximately 10 to 30 percent or more. Therefore, data gained by these methods have to be corrected in order to meet the data quality objectives given by the first European Air Quality Daughter Directive (expanded uncertainty of 25 % for 24 hour values).

In the ambient air monitoring network LUQS in North Rhine-Westphalia, PM10 measurements are performed at about 70 stations. At 30 stations gravimetric High Volume Samplers (Digitel DHA-80) were operated in parallel to continuous monitoring methods (TEOM SES and β -gauge FH 62 IR with temperatur control system TCS) in 2005. The evaluation of data gathered in previous years clearly shows that the gravimetric High Volume method is equivalent to the reference method described in EN 12341 (Low Volume Sampler). The expanded uncertainty at the limit value of 50 µg/m³ is lower than 12 percent. These results are documented in reports of the German ambient air monitoring networks.

In a simple and pragmatic approach, general correction factors of 1.26 ± 0.06 for TEOM SES and 1.14 ± 0.03 for FH 62 IR with TCS were calculated (2005). These corrections apply for the majority of stations. In some cases site-specific corrections are necessary. For these corrected PM10 data expanded uncertainties were evaluated using the procedure laid down in the European Guidance Document 'Demonstration of Equivalence of Ambient Air Monitoring Methods'. For 90 percent of all stations with parallel measurements the data quality objective of 25 % is met after correction. The averaged expanded uncertainty of the corrected FH 62 IR data is 14.3 % and lower than that for the TEOM SES (22.2 %, including cases exceeding the data quality objective).

The data analysis shows that simple correction factors for continuous PM monitoring methods can be easily derived from parallel measurements of continuous methods with a gravimetric reference method and work very well for annual averages, but are not ideal for the correction of daily averages. However, it is explicitly stressed that in spite of these results correction factors are indispensable and shall be applied anyway. Against the background of these results new developments in the field of continuous PM monitoring methods are watched carefully.

In North Rhine-Westphalia, the TEOM FDMS (Filter Dynamics Measurement System) was tested by LANUV (formerly: LUA) for PM10 following the requirements of the EU-guideline mentioned above. At a traffic-related site and an industry-related sit in Duisburg two TEOM FDMS (candidate method, CM) were operated over more than 6 months in parallel to two Digitel High Volume Samplers as the reference method RM (202 and 173 daily data pairs). The results of this equivalence tests are very promising: Between sampler uncertainties are calculated for the gravimetric method and the TEOM FDMS, respectively. The requirements of the guideline for this parameter (2 µg/m³ and 3 µg/m³) are safely met. No outliers were removed from the whole data sets. By orthogonal regression the following functions were evaluated: Traffic site (VDUI): CM = 0.96 * RM – 0.51; Industry site (DUBR): CM = 1.00 * RM – 2.54

The expanded uncertainty on a 95 % level of confidence, calculated according the EU guidance document without any correction, is 13.9 % (VDUI) and 17.1 % respectively (DUBR) at the limit value (50 µg/m³) and therefore well below the data quality objective of 25 %.

Additional campaigns for equivalence tests were performed by TÜV Rheinland (Cologne) for other types of instruments, for example the Thermo Model 5030 SHARP Monitor. This monitor was tested

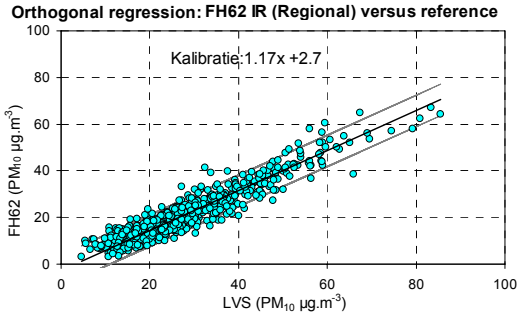
for PM_{10} and $PM_{2.5}$ against Leckel LVS3 or SEQ47/50 samplers at four sites according to the equivalence guideline. The requirements were met at all sites for PM_{10} and $PM_{2.5}$. Expanded uncertainties (without correction) for PM_{10} are in the range of 9 % to 14 %, for $PM_{2.5}$ in the range of 20 % to 23 %.

EQUIVALENCE OF PM10 MEASUREMENTS IN THE NETHERLANDS NATIONAL AIR QUALITY MONITORING NETWORK

Ronald Hoogerbrugge and Ruben Beijk
Dutch National Institute for Public Health and the Environment Email: Ronald.Hoogerbrugge@rivm.nl, Phone: 0031-(0)302743306

In the Dutch National Air Quality Monitoring Network (LML) particulate matter (PM10) is continuously measured at thirty-nine locations across the Netherlands. The automatic measurements are performed using two types of bèta-gauging (FH 62) monitors [1]. In total, 1646 validated parallel reference measurements at sixteen different measuring sites have been included in the equivalence study. The data are used for demonstration of equivalence between the automatic and reference method conform the guideline [2] as recommended by the Clean Air For Europe (CAFE) steering group.
The equivalence study shows significant differences in calibration functions between several monitor configurations and monitoring sites. Each is defined as a candidate measurement method. After application of the defined calibration the resulting relative (combined) measurement uncertainties vary between sixteen and twenty-two percent for the different candidates. Using the results all data, also historic data, are recalculated using the calibration factors and the application of the interim EU calibration factor of 1.3 is no longer necessary.

The statistical relation between the bèta-gauging monitors and the reference method is based on orthogonal regression. In case of an insignificant intercept the regression is forced trough the origin. The guidance document does not recommend the orthogonal regression through the origin due to the lack of an algebraic expression for the associated uncertainty. Therefore an algebraically method is derived. The method is validated by the statistical boots trap approach.



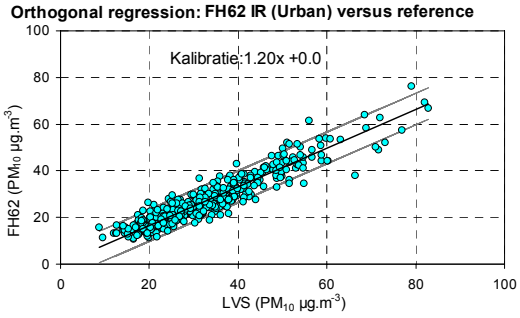


Figure 1: Orthogonal regression with the FH62 measurements at regional sites concentration in the Netherlands in the period 1994-2006 after applying the calibration as measurements on the x-axis. Above and under determined in the 2006 equivalence study. the regression the 95%-confidence borders. Distinction is made between regional, citybackground and street sites.

Figure 2: Development of the year average PM10 concentration in the Netherlands in the period on the vertical axis versus the parallel reference measurements on the x-axis. Distinction is made between regional, citybackground and street sites.

[1] Beijk, R., Hoogerbrugge, R., Hafkenscheid, T.L., Arkel, F.Th. van, Stefess, G.C., Meulen, A. van der, Wesseling, J.P., Sauter, F.J., Albers, R.A.W. (2007) PM10: Validatie en equivalentie 2006. Bilthoven: National Institute for Public Health and the Environment (RIVM). Publication 680708001

[2] EC Working Group on Guidance for the Demonstration of Equivalence (2005) Demonstration of equivalence of ambient air monitoring methods. Published and available on the internet: http://ec.europa.eu/environment/air/ambient.htm (February 2007)

PRACTICAL TESTING OF EQUIVALENCE DEMONSTRATION FOR PM AUTOMATED MONITORING METHODS: THE EXPERIENCE IN FRANCE

O. Le Bihan[1] , F. Mathé[2], H. Marfaing[3], D. Robin[4]
1) LCSQA-INERIS, Parc technologique Alata - 60550 Verneuil-en-Halatte, France (Olivier.Le-Bihan@ineris.fr). 2) LCSQA-Mines de Douai, 59508 Douai, France (mathe@ensm-douai.fr). 3) Airparif, 7 rue Crillon, 75004 Paris, France (helene.marfaing@airparif.asso.fr). 4) Airmaraix, 67-69 Avenue du Prado 13286 Marseille, France (dominique.robin@atmopaca.org).

Introduction

In France, measurement of the mass of particulate material in the atmosphere (mainly as PM10, but including some $PM_{2.5}$ monitoring) is carried out by 36 air quality monitoring networks with nearly 450 SPM automatic monitoring stations. These networks are operated at both the regional and local levels, on behalf of the french state, the local authorities, industries and environmental organisations. Since the 90's, automatic monitoring of particulate matter has been largely founded on the TEOM and beta attenuation analysers (85% vs 15%).

The present study summarizes the results of measurement campaigns conducted in the framework of LCSQA activities. The main objective is to show that the methods tested (TEOM – FDMS from Thermo R&P and beta gauge MP101M-RST from Environnement SA) meet the Data Quality Objectives for fixed measurements specified in the Air Quality Directive [1], under conditions reflecting practical application in air quality monitoring networks. The principles and methodologies to be used for checking the equivalence of non-reference methods for PM-monitoring are given in a document written by European Commission Working Group [2].

Method

In each case, the experimental sites are located in an urban background or suburban area and near an air pollution monitoring station from an air quality network. The choice of site and time period of the year has been based on representativeness for typical conditions for which equivalence will be claimed, including possible episodes of high concentrations. All tests have been performed in which all methods are compared side-by-side. Concerning particulate matter PM_{10} and $PM_{2.5}$, the gravimetric reference used for french campaigns was a Partisol Plus 2025 sequential sampler (Thermo R&P) collecting on 47 mm diameter, 2 µm pore size PTFE filters (ZefluorTM from Pall Corporation).

Concerning fields tests performed in Belgium and Italy [3], only the beta attenuation monitor has been studied for the PM_{10} size fraction. The gravimetric reference used was respectively a Leckel SEQ 47/50 sequential sampler (Thermo R&P) collecting on 47 mm diameter, Quartz Fiber filters (from Macherey-Nagel) and a Reference Sampler designed by CNR-IIA collecting on 47 mm diameter, 2 µm pore size PTFE filters (TefloTM from Pall Corporation).

Concerning Bobigny and Marseille trials, the weighing procedure was conducted by INERIS according to requirements of European Standard EN 14907 [4], using a balance with a resolution of 10 µg in a temperature (20 ± 1°C) and humidity (50 ± 5% RH) controlled weighing room. For trials abroad, the weighing operations were performed respectively by VMM and CNR-IIA.

All apparatus were duplicated and equipped with same size-selective inlet as the reference method. Sampling and monitoring series have had a duration sufficient to collect a minimum of 40 duplicated pairs of measurement results each averaged over at least 24-hour per comparison.

Results and discussion

The following tables provide an overall summary of results of each trials for each instrument included (Thermo R&P FDMS series 8500 version b for PM_{10} and $PM_{2.5}$, Environnement SA MP101M-RST for PM10).
The criteria of equivalence demonstration procedure are:
- datasets suitability (at least 20% of the results are greater than 25 µg.m-3 which is 50% of the daily limit value specified in the first Daughter Directive)
-between-instrument uncertainty (3 $\mu g.m^{-3}$ not to be exceeded),
- slope and intercept orthogonal regression line equation must be non significative according to recommended statistical data treatment recommended
-expanded relative uncertainty at the level of daily limit value (50 $\mu g.m^{-3}$ for PM_{10}, assumed to 25 $\mu g.m^{-3}$ for $PM_{2.5}$) observed during the test must meet data quality objective of ±25 % required by Directive

Candidate method	Trial site location	Time period	Equivalence criteria met?
Thermo R&P TEOM-FDMS PM_{10} series 8500 version b (a)	Bobigny (France)	25/01 to 17/04/2005	Yes
	Marseille (France	21/12/2005 to 13/04/2006	
Thermo R&P TEOM-FDMS $PM_{2.5}$ series 8500 version b (a)	Bobigny (France)	25/01 to 17/04/2005	Yes
	Marseille (France	21/12/2005 to 13/04/2006	
Environnement SA MP101M-RST PM_{10} (b)	Bobigny (France)	25/01 to 17/04/2005	Yes
	Marseille (France	21/12/2005 to 13/04/2006	
	Aarschot (Belgium)	10/05 to 24/06/2006	
	Monterotondo (Italy)	24/06 to 19/08/2003 & 10/12/2003 to 11/01/2004	

(a) 1h-step time measurement
(b) 24h-step time measurement

Results for the two automatic monitoring methods show that the following meet the equivalence criteria set out: TEOM retrofitted with FDMS (for PM_{10} and $PM_{2.5}$); and beta gauge MP101M-RST (for PM_{10}). All these units meet the equivalence criteria set down without the application of correction for slope and/or intercept. Due to the variability of test sites (in time and space) involving different composition of ambient air and meteorological conditions, it can be assumed that equivalence for equipment tested under the used configuration is valid anywhere else in France under ambient conditions. The authors acknowledge the manufacturers and the Belgian (VMM – ISSeP – IBGE BIM) & Italian (CNR-IIA) institutions for their technical support. This work was supported by the French Ministry of Environment (MEDD) and by the French Agency for Environment and Energy Management (ADEME) in the framework of Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) activities

Keywords: Demonstration of equivalence, Data Quality Objective, EU Directives, automatic monitors.

[1] Council Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air

[2] « Demonstration of equivalence of ambient air monitoring methods » - Report by an EC Working group on Guidance for the Demonstration of Equivalence (November 2005)

[3] Technical Report « Evaluation of automatic analyzer MP101M.C for mass concentration measurement of PM10 particulate matter upon request by Environnement SA using reference procedures required by Ministerial Decree n°60 – 2 april 2002 – in application of Directives 1999/30/CE – 2000/69/CE » -17/11/2004

[4] Standard EN 14907 (2005) « Ambient air quality - Standard gravimetric measurement method for the determination of the $PM_{2.5}$ mass fraction of suspended particulate matter »

EQUIVALENCE OF DIFFUSIVE SAMPLERS TO REFERENCE METHODS: THE MEMBRANE-CLOSED PALMES TUBE AND OZONE RADIAL DIFFUSIVE SAMPLER EXAMPLES

Michel Gerboles and Daniela Buzica
DG-JRC, Institute for Environment and Sustainability, Joint Research Centre Via E. Fermi, I-21020 Ispra VA, Italy. michel.gerboles@jrc.it Tel: +39 0332 785652, fax: +39332 789931.

The study of the equivalence of the membrane-closed Palmes tube (MCPT) to the chemiluminescence method, the European Nitrogen Dioxide (NO2) reference monitoring method, is presented. In the MCPT, a membrane is added at the open end of the Palmes tube to avoid the effect of turbulence due to wind speed. The MCPT is further optimized for the volume of absorbent and number of meshes retaining the absorbent. Finally, the MCPT is placed in an air proof container and is ready to use in order to improve the quality of measurement. Experiments were carried out in an exposure chamber in order to establish an equation able to predict the uptake rate of the MCPT according to the conditions of exposure. By studying the diffusion along the tube path, a formula that only depends on the logarithm of the water vapour concentration was established. The empirical equation showed a significant effect arising from temperature and humidity. Small effects arising from wind speed, NO2 and averaging time (provided that this period exceeds 3 days) were also observed. The equation was then applied on some field data and laboratory inter-comparison tests. The agreement between the chemiluminescence method and diffusive samplers changes from 20% of mean bias with the traditional Palmes diffusion tubes to 5% of mean bias with the membrane-closed Palmes tube (MCPT) at the same monitoring site. Further to the requirements of the guide for demonstration of equivalence, the MCPT has shown to be able to correctly measure NO2 under two extreme exposure conditions. Open Palmes diffusion tubes were reported to produce biased measurements. On the contrary, the MCPT was shown to give good results under a wide range of conditions. By calculating the uncertainty according to the guide for the demonstration of equivalence, the data quality objective (DQO) of the first Daughter Directive, accuracy lower than 25% for indicative method was met with the MCPT. Moreover, the MCPT satisfied the DQO of 15% for continuous measurement at the annual limit values of 40 µg/m³.

PARTICULATE MATTER EQUIVALENCE IN THE UNITED KINGDOM

Janet Dixon
Department for Environment, Food and Rural Affairs, U.K.
janet.dixon@defra.gsi.gov.uk, Tel. +44(0)2070 828372, Fax +44(0)2070 828379

In the UK, measurements of PM10 are largely founded on the use of the Tapered Element Oscillating Microbalance (TEOM). This paper summarises the results of the UK Equivalence Programme for Monitoring Particulate Matter; a cross-comparison between candidate PM_{10} and $PM_{2.5}$ sampling methods with the PM_{10} reference method (EN12341) in the first Daughter Directive (1999/30/EC) and $PM_{2.5}$ reference method (EN14907) to show equivalence or otherwise.

The UK programme included the operation of seven candidate instruments collocated with the EU reference method (Low Volume Samplers (Kleinfiltergerat) for PM_{10} and $PM_{2.5}$). Instruments included are: Tapered Element Oscillating Micro-balance (TEOM); TEOM retrofitted with Filter Dynamics Measurement System (FDMS; for both PM_{10} and PM2.5); Partisol 2025 Sequential Sampler; OPSIS SM200 (Beta and Mass configurations) and Met-One Beta Attenuation Monitor (BAM).

The programme included operation of the monitoring equipment across eight studies that collectively represent particulate conditions typically encountered across the UK. Field campaigns have been undertaken at four locations during summer and winter periods. All instruments were operated in duplicate in order to determine the 'between-sampler' uncertainty.

In applying the criteria contained within the Guidance for Demonstration of Equivalence (the Guidance) to the UK programme, it has been necessary to make some interpretation of the current criteria:

- corrections have been applied where the slope of a data set was either all greater or all less than 1 and/or where the intercepts of the data set were either all greater or all less than 0;
- if a candidate instrument failed on WCM for the <50% LV dataset alone (either before or after slope and/or intercept correction), this was not considered sufficient evidence for a candidate instrument to be excluded. The results of the programme show that the current TEOM monitoring method used in the UK fails to meet the criteria for equivalence set out in this study. This result is consistent with previous investigations reported by Defra and the devolved administrations.

Further work has been carried out to derive, test and apply a model to correct TEOM PM10 measurements such that they can be used to measure PM10 for assessment against the EU Limit Value in the UK. The KCL Volatile Correction Model (VCM) was based on analysis of daily mean measurements of PM10 by Filter Dynamics Measurement System (FDMS) and TEOM at sites in the UK Equivalence Programme and at sites in the London Air Quality Network (LAQN). The model used the FDMS purge measurement (a measure of the volatile component of PM_{10}, which is expressed as a negative concentration) to correct for differences in the sensitivity to volatile PM_{10} between the TEOM and the EU PM_{10} reference method. The model equation for the correction of TEOM PM_{10} measurements is:

$$\textit{Reference Equivalent } PM_{10} = \textit{TEOM} - 1.87\ \textit{FDMS purge}$$

The FDMS purge concentration may be measured at a remote site, allowing the possibility of using a single FDMS instrument to correct PM_{10} measurements made by several TEOMs in a defined geographical area. To test the equivalence of the model to the EU PM_{10} reference method, the model was assessed against the criteria within the Guidance. Overall 772 equivalence tests were undertaken.

The model passed the Guidance equivalence criteria at the sites used in the UK Equivalence Programme and can therefore be considered an equivalent method. Further, the model passed the equivalence criteria using remote FDMS purge measurements over a maximum distance of approximately 200 km. This shows the model to be a viable tool for correcting measurements from TEOM instruments on the national and local government networks using FDMS purge measurements from a more limited network of sites.

Results for other instruments in the Equivalence Programme show that the Partisol 2025 Sequential Sampler; TEOM retrofitted with FDMS (for PM_{10} and $PM_{2.5}$); and the OPSIS SM200 (by Beta) meet the equivalence criteria set out in this study without the application of correction for slope and/or intercept. The OPSIS SM200 (by Mass) and Met One BAM meet the equivalence criteria set down in this study only after application of correction factors for slope and/or intercept. The operation of candidate instruments in configurations different from those employed in this study may constitute a different method, and it cannot be assumed that the conclusions are transferable. Further field work is currently ongoing to test the new FDMS model and the $PM_{2.5}$ Met One BAM.

PM10 EQUIVALENCE TESTS IN SWITZERLAND: RELATIONSHIP BETWEEN A TEOMFDMS AND THE REFERENCE METHOD AT A SUBURBAN SITE

Ch. Hueglin, R. Gehrig
Empa, Swiss Federal Laboratories for Materials Testing and Research, CH-8600 Duebendorf, christoph.hueglin@empa.ch, Tel. +41 44 824 46 54

Automated PM10 monitors (TEOM-FDMS 8500, Thermo Electron Corporation, East Greenbush NY, USA) were tested in the field for equivalence with a manual gravimetric method fulfilling the equivalence criteria of EN12341 (PM10 collection using a high volume sampler HVS, Digitel DA-80H). Two TEOM¬FDMS instruments were run collocated at the suburban site in Duebendorf (Switzerland) from August 26 2005 to December 11 2006. The data were recorded on a hourly basis and aggregated to daily values. During the same measurement period, a HVS equipped with glass fibre filters (HVS/GF, Ederol 227/1/60) was operated at a total of n=102 days. In addition, a second HVS equipped with quartz fibre filter (HVS/QF, Whatman QMA) was run on n=109 days during the January 1 2006 to December 11 2006 period.

The between sampler uncertainty of the candidate method (CM – here the TEOM FDMS) was calculated from a total of 350 parallel 24-hour PM10 measurements, and determined to 1.64 23 µg/m3. This value is somewhat larger than the corresponding value for the reference method (RM). Based on 31 parallel 24h¬PM10 measurements from September 27 2005 to November 26 2005, the obtained between sampler uncertainty of the RM is 0.92 µg/m3.

The comparison of the TEOM FDMS and the HVS equipped with glass fibre filters was excellent, a orthogonal regression analysis yielded a relationship between the CM and the RMHVS/GF of CM=1.01(± 0.01)•RMHVS/GF -0.56(± 0.55). The daily PM10 values used for this analysis ranged from 4.2 to 181.0 µg/m3. Exceptionally high PM10 levels occurred in January and the beginning of February 2006, when stable high-pressure weather was leading to temperature inversions and accumulation of air pollutants within the boundary layer in large parts of Switzerland and the neighbouring countries. During this high air pollution episode, the HVS equipped with quartz fibre filters showed substantial losses of PM10. This can be concluded from the fact that during this time period the PM10 values measured with the HVS/QF were systematically lower than with the HVS/GF and the two TEOM FDMS instruments. Consequently, the slope of the linear relationship between the CM and RMHVS/QF as estimated by orthogonal regression was significantly larger than one, CM=1.11(± 0.01)•RMHVS/QF - 1.54(± 0.52).

Nevertheless, the relationship between CM and RMHVS/QF is equally convincing than for the CM RMHVS/GF comparison when measurements during the high air pollution episode are rejected (CM=1.03(± 0.01)•RMHVS/QF - 0.22(± 0.31), n=94, the range of RMHVS/QF PM10 is 3.9-58.3 µg/m3).

In conclusion, the PM10 field measurements at an urban background site in Switzerland during all seasons indicate that TEOM FDMS monitors can be accepted as equivalent to the reference method. It is on the other hand worrying, that for the gravimetric method using HVS with QF, PM10 losses lead during high concentration episodes to negatively biased PM10 concentration measurements

FIELD CAMPAIGNS TO DETERMINE THE CORRECTION FACTORS IN THE SPANISH ATMOSPHERIC NETWORKS

García Dos Santos S. , Díez Hernández P. and Fernández Patier
R. Área de Contaminación Atmosférica. Centro Nacional de Sanidad Ambiental (ISCIII). Ctra. Majadahonda-Pozuelo km 2, 28220 Majadahonda (España)

Introduction

In 1999, The Spanish Reference Laboratory for Atmospheric Pollution (ACA), obtained its accreditation as an essay laboratory, according to the requirements of the EN ISO/IEC 17025 by ENAC (Spanish Body for Accreditation) EN 12341: 1998 standard. The ACA has been performing the calculation of the default correction factors between the reference method (EN 12341: 1998 standard) and the analysers set by the networks to measure PM10 concentrations. This work started in June 2003 and over last until today.

Material and Methods

In order to calculate the correction factors, the ACA used the requirements set on chapter 5 of the EC GUIDE (Guidance to Member States on PM10 Monitoring and Intercomparisons with the Reference Method) from 22 January 2002. Strictly speaking the GUIDE requires the following:

¾ There should be at least 30 days continuous data measurement in any given period (summer or winter)
¾ Data shall be compared with EN 12341 standard or a proven equivalent instrument
¾ Only regression equations having $r2 \geq 0.8$
¾ Data with a regression equation constant > 5 μg/m3 (b from y = ax + b) in absolute values shall be rejected

Therefore, the ACA performed 29 field campaigns (winter and summer) in 7 Spanish networks including 14 field stations. Also, the types of analysers of the stations were varied from B-Attenuation, to Optical or Tapered Oscilatory Microbalance including different brands and models.

Results

The results of the campaigns are showed on **Table 1**. From the table is clear that many networks do not meet the Guide requirements for either r2 or b or both. Also, there are differences of the calculated default factors. Moreover, some factors are very below 1,00 (≈ 0,73) which means the station is given PM10 concentrations higher than the ambient levels. However, other stations showed values closed or above 1,00. Notice, main of these last values have been obtained during the 2006 and onwards campaigns. This suggested an improvement in the quality of the data obtained by the networks. Also, clearly some networks (nº 6 and nº 7) have showed a better quality than others, perhaps due to have better quality systems.

From the results, is possible to say that some analyser types did not measure PM10 particles quite well. But, nowadays it is unclear to know if this is because of bad network systems or to the analyser performance. However, with theses equivalence tests some networks have improved their commitment to obtain more reliable results.

Finally, the ACA is going to recalculate the results using the procedure set by the new Equivalence Guide published on November 2005.

Table 1: Field campaigns to determine the correction factors in the spanish atmospheric networks

Network	Field Stations	PM10 Type Analyser	Data of field campaigns	Results
1	1 A	β-Atenuation	Summer : 08.07.2003 - 27.08.2003	Factor: 0,73
	1 B	β-Atenuation	Summer: 08.07.2003- 27.08.2003	Not met GUÍDE for b
	1 A	β-Atenuation	Winter: 24.02.2004 - 14.04.2004	Factor: 0,91
	1 B	β-Atenuation	Winter: 24.02.2004 -14.04.2004	Not met GUÍDE for b and r_2
	1 A	β-Atenuation	Summer: 27.09.2004 -10.10.2004	Not met GUÍDE for b
	1 B	β-Atenuation	Summer: 27.09.2004-10.10.2004	Not met GUÍDE for b
2	2 A	β-Atenuation	Winter: 06.11.2003-16.12.2003	Not met GUÍDE for r_2
	2 B	β-Atenuation	Winter: 06.11.2003- 16.12.2003	Not met GUÍDE for b and r_2
	2 A	β-Atenuation	Summer: 21.06.2004 -25.07.2004	Factor: 0,74
	2 B	β-Atenuation	Summer: 21.06.2004 -25.07.2004	Not met GUÍDE for b
3	3 A	β-Atenuation	Summer: 11.06.2005 -04.10.2005	Not met GUÍDE for b and r_2
	3 B	Laser	Summer: 11.06.2005 -04.10.2005	Not met GUÍDE for r_2
	3 A	β-Atenuation	Winter: 29.12.2005 -17.12.2005	Not met GUÍDE for b and r_2
	3 B	Laser	Winter: 29.12.2005 -17.12.2005	Not met GUÍDE for r_2
4	4 A	β-Atenuation	Summer: 15.08.2005 - 13.06.2005	Not met GUÍDE for b and r_2
	4 A		Winter: Campaign not performed	
	4 B	β-Atenuation	Winter: 24.03.2006 - 08.02.2006	Not met GUÍDE for b
	4 B	β-Atenuation	Summer: 18.11.2006 - 28.08.2006	Factor: 1,22
5	5 A	β-Atenuation	Winter: 20.03.2006 - 26.01.2006	Factor: 0,88
	5 B	β-Atenuation	Winter: 20.03.2006 - 26.01.2006	Not met GUÍDE for b and r_2
	5 A	β-Atenuation	Summer: 26.09.2006 - 31.07.2006	Not met GUÍDE for b and r_2
	5 B	β-Atenuation	Summer: 26.09.2006 - 31.07.2006	Factor: 0,97

Network	Field Stations	PM10 Type Analyser	Data of field Campaigns	Results
6	6 A	TEOM	Summer: 06.08.2006 - 31.12.2006	Factor : 0,99
	6 B	TEOM	Summer: 06.08.2006 - 31.12.2006	Factor: 1,00
	6 A	TEOM	Winter: In campaign now	
	6 B	TEOM	Winter: In campaign now	
7	7 A	β-Atenuation	Summer: 28.07.2006 - 19.10.2006	Factor: 1,02
	7 B	TEOM	Summer: 28.07.2006 - 19.10.2006	Factor: 0,98
	7 A	β-Atenuation	Winter: In campaign now	
	7 B	TEOM	Winter: In campaign now	

DEMONSTRATION OF EQUIVALENCE FOR PM MONITORS IN FLANDERS, (BELGIUM)

J. Vercauteren, C. Matheeussen, E. Roekens
Flemish Environment Agency, Kronenburgstraat 45, 2000 Antwerp, BELGIUM
e-mails: j.vercauteren@vmm.be, c.matheeussen@vmm.be , e.roekens@vmm.be
Tel. +32 (0)3 2441232

Of the three Belgian agencies that monitor PM in ambient air the Flemish Environmental Agency (VMM – Vlaamse Milieumaatschappij) is responsible for Flanders. General monitoring is carried out by means of the 'telemetric' network which uses beta attenuation (FH62-IR) monitors. Measurements at special and hotspot locations is done by the 'specific studies' network which uses both normal TEOM (1400a) instruments and TEOM-FDMS systems. In addition to these 3 types one laser-based system (Grimm #180) is in test. For PM equivalence demonstration and determination of heavy metals in ambient air Leckel SEQ-47 instruments are used as 'reference method'. Although the demonstration of equivalence has been restricted to PM10 up to 2006, the more stringent PM2,5 guidelines concerning filter weighing have been in use for several years.

Currently Flanders still uses some of the highest correction factors for PM10 in Europe: 1.37 for the beta attenuation monitors and 1.47 for the normal TEOMs. These two factors were calculated in 2002 after comparative campaigns in which cellulose nitrate was used as reference filter. At the time cellulose filters showed better characteristics for metal determination, were found equal with quartz filters during a preceding campaign and little was known about the effect of the filter material on the collection of PM.

In 2005 and 2006 new comparison campaigns were carried out, this time not only replacing cellulose filters by the prescribed quartz fibre but most of the time even using 2 different brands of quartz fibre, namely Macherey-Nagel (QF10) and Whatman (QM-A). QF10 was selected because it gave better blank values for metal determination, QM-A because it is widely used in other parts of Europe and the US. The remarkable outcome of the tests was that the PM10 concentrations on QM-A filters were on average around 15% higher than on QF10 (for one winter campaign the difference was almost 25%). For both filter brands a capture efficiency test with a typical test aerosol (oil droplets) gave results better than 99,9%. The reason for the higher concentrations on QM-A is believed to be a different efficiency for semi-volatile compounds like ammonium nitrate and/or water. Other things that came out of these campaigns were significant differences between reference samplers inside and outside shelters and between pre-fired quartz fibre filters and not pre-fired filters, the latter indicating that 48 hours of acclimatisation at 50% RH was not sufficient.

The results of the 2005-2006 campaigns allowed VMM to calculate double correction factors: for QF10 and for QM-A as reference. For the former they turned out to be lower than the currently used factor, for the latter they were close to the existing factors. TEOM-FDMS turned out to be equivalent, without correction, with QF10. But for QM-A a factor of around 1.1 was required. The expanded uncertainty criterion only turned out to be a problem for the regular TEOMs. Since the current European standard does not favour one of these two filter brands over the other VMM has decided to stick to the existing factors (and not to use a factor for TEOM-FDMS) for the time being.

At this moment VMM is carrying out a large, yearlong PM10 chemical characterisation project in which quartz fibre (this time only QM-A) and Teflon membrane filters are being used for determination of soluble ions, EC/OC and elements. The filter masses are also used for comparison with automatic monitors and the chemical information should allow for a better understanding of the variations in differences between automatic monitors and reference samplers. On top of this comparison campaigns for PM2,5 monitors are also underway.

Many observations in VMM's story point to the same conclusion: the issue of equivalence is highly interconnected with the improvement of the European PM10 (and PM2,5) reference method which is currently high on TC264 WG15's agenda. A new reference method should specify it's working conditions in a very precise and strict way so that application of it really gives the same results all over Europe. The best way to guarantee this seems to be a choice for the scientifically best method. A second key to reach equivalent results all over Europe is a uniform procedure for the demonstration of equivalence.

QUALIFYING A PM10/PM2,5 METHOD AS TRANSFER STANDARD OR FOR AUTOMATED STANDARD: EXPERIENCE WITH THE SWAM5-A

L. Bertrand, G.Gérard and S.Fays
Institut Scientifique de Service public (ISSeP), Liège, Belgium
e-mail : L.bertrand@issep.be

As other ambient air monitoring networks, the Walloon network in Belgium faces a number of difficulties to establish timely Equivalence of PM monitoring equipment used.
For historical reasons, the instrument used in the telemetric network is the beta attenuation analyser MP101M.C of Environment-sa. The "C" stands for the upgrade with a Regulated Sampling Tube system. That improvement saved the device, in spite of a marked lack of sensitivity. In the Walloon network, the MP101M.C is run with a cycle of 2H, 6H or 24H, depending on site.

In a 2006 collaborative trial in Aarschot, this beta attenuation analyser succeeded well at Equivalence tests when run with a 24H cycle. Possible correction factor (1.00 to 1.19) depended on quartz filter brand chosen as gravimetric reference. With a 6H cycle, daily values were 78% of that with a 24H cycle. With that 6H cycle; and a single set of data, after an intercept correction of around 6 to 8 µg/m³ (according to quartz reference used), the outcome was Equivalence. As designated air reference Lab, we were requested to try further using that MP101M.C with the shorter cycles, (and get legalizable daily values). This increases significantly the workload of necessary validation trials which ISSeP has to cope with.

ISSeP has been using and testing two dual channel samplers and beta attenuation analysers Swam5-A of FAI-Instruments (Fonte-Nuova, Italy) since October 2006. We experienced its outstanding reliability and user-friendliness. Precision of flow was better than what ISSeP ever experienced before (0,7% - geometric mean of relative deviations, measured independently, from nominal flow), and it's the same for analyser precision (u.bs of 0,63 to 0,69 µg/m³ using fibreglass filters). Initial results with quartz filters (before remedies/trouble shooting) were analyser u.bs<2 µg/m³ and beta analyser against gravimetry ratios of 1.00 (PM10) and 0.98 (PM2,5).

Continuous instrument recording of temperatures (ambient air and filter vicinity) shows how long during the 24H and to what extent that the standard criteria (Difference < 5°C) is not fully met (will it be some day by any commercially available device ?)
Remedies of Swam manufacturer in reaction to ISSeP's feedback as a user were an upgrade of the filter cartridge (to avoid filter adhesions observed with some brands of quartz filters), a subsequent mechanical adjustment to maintain perfect shape of dust spots submitted to the beta attenuation analysis and an offset determination experience (followed by derived embarked offset correction). New results with quartz filters (40+days) in thus optimised conditions are due mid-April 07.

Pending decisions for ISSeP as a designated Lab and as a monitoring network relate to questions which all designated Reference Labs and most member states face presently:
Does it make sense to consider comparison of a candidate analyser with gravimetry as more appropriate than comparison with a a high precision sophisticated automatic indirect measurement (by beta attenuation) performed on site within a few hours after completion of sampling, and precisely on a standard compliant filter sampling of 24H ? (For practical reasons gravimetric comparisons are run with 14 days batches of filters and no brand of quartz filter nor lab personnel is totally immune to filters handling damages/precision losses, as gravimetric u.bs generally show). If a Transfer (Automatic standard) method is used,- in the Walloon case with far more extensive comparison (than the 4 x 40 days requirement of Equivalence document) of Walloon network analyser against Swam5a beta analyser considered as a Transfer standard- , how will EC deal with the resulting Equivalence demonstration/report trials – which are to legalize network PM data?

EQUIVALENCE OF PM10 IN SLOVENIA

T.Bolte, P.Pavli.
Environmental Agency of the Republic of Slovenia,
e-mail: tanja.bolte@gov.si, Tel.+386 478 4036, Fax.+386 478 4050

Air pollution adversely affects the health of humans and other living beings. Besides industry and fuelling, traffic is the major source of air pollution with particulate matter. But we have a problem of losses of volatile organic compound in TEOM continuous monitoring, which is performed at seven stations in Slovenia, so we must introduce a correction factor determined from the reference monitoring of particulate matter with Leckel instruments. Correction factor differs from site to site and is seasonly dependent (winter/summer).

Monitoring of particulate matter is performed in national network managed by Environmental Agency of the Republic of Slovenia (EARS) and in the frame of supplementary networks, managed and financed by local communities of larger cities (i.e. Ljubljana, Maribor, Celje) or by some factories, which are large sources of emmissions (e.g. thermo-power plants). In national network, which was modernized in 2001, continuous measurements of PM10 are performed by TEOM (with PM10 inlet) at 7 different locations (Ljubljana, Celje – urban background); Maribor, Zagorje, Trbovlje – urban traffic; Murska Sobota – rural background; Nova Gorica – suburban background). Continous measurements of PM10 are performed by TEOM. As we know, the inlet and filters of these monitos are generally heated to 35-50 deg. C. At this temperature volatile particles (like ammonium nitrate) are partly lost. Therefore these data are not directly comparable to manual gravimetric data. We started with PM10 measurements with Leckel instrument, according to SIT EN 12341. This is gravimetric method. Samplers are collected on filters, during 24-hours period. The choice of filter type is dependent on instrumenmtation and on what type of analysis is going to be done after sampling. We normally used glass of quartz filters.

The intercomparison measuremets between two samplers are performed according to Guidance to memebr states on PM10 monitoring and intercomarisons with the reference method. The candidate samplers were installed on top of measuring containers. Due to the height of containers of circa 2.5 m, the sampling inlets of various instruments were at a height of circa 3.5 m above ground level. Minimum distance between sampler inlets are 70 cm. The flow rates of low volume samplers were checked every two weeks with rotameter and Dry Cal, and readjusted if necessary. For each measuring site we defined two correction factors, one for summer period (from 01.04. to 30.09.) and one for winter period (from 1.10. to 31.03.). Correction factors differ from site to site. **Table 1** shows correction factors at measuring sites in Slovenia in year 2004. For the future work, we will make the evaluation of data from both samplers, and determine the uncertainty of these measurements.

Table 1: Correction factors at measuring sites in Slovenia.

Measuring site	Type of station	Period	Correction factor
Ljubljana Bežigrad	Urban background	Winter /Summer	1,24/1,03
Maribor	Urban traffic	Winter /Summer	1,19/1,00
Celje	Urban background	Winter /Summer	1,12/1,30
Murska Sobota	Rural background	Winter /Summer	1,22/1,10
Nova Gorica	Suburban background	Winter /Summer	1,20/1,30
Trbovlje	Urban traffic	Winter /Summer	1,30/1,30
Zagorje	Urban traffic	Winter /Summer	1,39/1,30

NON SEASONABLE VARIABILITY AND THE EFFECT OF LOGGING AVERAGES ON TEOM AND FDMS MEASURMENTS MADE DURING THE UNITED KINGDOM PM_{10} and $PM_{2.5}$ EQUIVALENCE TRIALS

David Harrison, Jeff Booker and Richard Maggs
Bureau Veritas HS&E Ltd., 30 Great Guildford Street, London SE1 0ES, United Kingdom.
Email: david.harrison@uk.bureauveritas.com ; Tel: 0207 9026170

The United Kingdom Equivalence trials took place from late 2004 to early 2006, and the primary results were published then presented at the meeting in Antwerp during June 2006. This presentation highlights two crucial sets of findings that were not covered in the report.

Previous studies in Switzerland and the United States have shown that the TEOM versus filter measurement correction factor varies with ambient temperature. The results of the UK equivalence study were reanalysed using ambient temperature as a variable. There was shown to be very little correlation, and a correction method for TEOMs incorporating ambient temperature would not be suitable for deployment in the United Kingdom. It is thought that the ambient temperature effects observed in the previous studies are masking the seasonal variability of volatile species such as ammonium nitrate. In the United Kingdom, ammonium nitrate concentrations are highly variable both geographically and temporally, and as such do not follow the seasonal variation of ambient temperature.

TEOMs in the United Kingdom National Network are typically logged using data loggers that operate on a 15 minute averaging time base. The effects on the data quality due to logging TEOMs in this manner are significant for 15 minute averages, but could be considered insignificant for hourly data. The effects of attaching an FDMS to the same logging system are discussed. It is shown that this method significantly effects the concentrations obtained from the instrument for hourly measurements. These effects are reduced for 24 hour averages; however, this method is still shown to be unsuitable for wider deployment. The potential implications of making FDMS measurements at time periods of less than 1 hour are discussed, along with details of how the second round of UK PM Equivalence Studies will investigate methods of logging FDMS data every minute in order to calculate true averages.

EXPERIENCE ON DEMONSTRATION OF EQUIVALENCE on PM_{10}/$PM_{2.5}$. 3 ISSUES.

D. de Jonge
Department of Environmental Medicine, Municipal Health Service Amsterdam, The Netherlands, ddjonge@ggd.amsterdam.nl, Tel. +31205555405, Fax +31205555690.

PM measurements

The air quality measurement network of Amsterdam counts currently 12 measuring stations, some are in operation from the early seventies. TEOM systems (operated at 50 °C) are applied since 1998 for the measurement of the PM10 concentrations on 5 locations, 3 locations for $PM_{2,5}$ and 4 locations for the PM_1 concentration.

At 2 locations the PM_{10} and $PM_{2,5}$ concentrations are measured every 4rd day since 2004 with the reference method according to respectively EN 12341 and EN 14907. At 1 location a PM2.5 TEOM FDMS is being tested.

Methods

To test the equivalence of the TEOM measurements the standard excel sheet provided by the EU [1] has been used. Other examples of (results) of the equivalence test have been investigated [2,3] These reports seem to have different views on some evident details. Calculations with our data sets have been done but, due to unclear rules, have not been finished. Which resulted in some discussable issues.

ISSUE 1: It is unclear if the CM is defined including or excluding corrections. If uncorrected data is used in the Excel input sheet, it will calculate the optimum correction in each case. This optimum correction is probably different for each location.

In practice most network use a standardised factor (f.e. 1.3) for the whole network. It is possible that for this standard correction the equivalence fails on individual locations.

ISSUE 2: The DEFRA report on the equivalence of PM measurements evaluates the candidate method for each single site and for all data from all sites together. The RIVM report aggregates data per type of location (up to 8 stations) to test the equivalence.

In the DEFRA case a more stringent judgment has been made. For example the DEFRA concludes that the TEOM “Does not meet equivalence criteria”. If the corrected TEOM data (page 89) are analysed analogue to the RIVM report, it does meet the criteria, although 4 out of 8 locations of the individual campaigns fail.

ISSUE 3: The size of the dataset with CM and RM to test the equivalence seems to be unlimited. Both the DEFRA report and RIVM report aggregates data from several years. Since the test is aimed to demonstrate equivalence for annual averages, it is questionable if datasets of several years may be aggregated to one test.

Discussion

The EU equivalence test should be prescriptive on subjects below:

- evaluation should be done on a corrected CM, if a correction in reality is used,
- evaluations should be made for separate sites individually to answer the question,

- if and where a CM is equivalent or not a dataset reference measurements should be taken in one year as maximum period, because equivalence of the annual average concentration is one of the items to evaluate, so the test sheet should be limited to a maximum of 365 data pairs.

[1] ec.europa.eu/environment/air/pdf/test_equivalencev31004.xls, build in October 2004.

[2] Harrison, D. UK Equivalence Programme for Monitoring of Particulate Matter. June 2006. www.airquality.co.uk/archive/reports/cat05/0606130952_UKPMEquivalence.pdf

[3] Beijk, R et al. PM10: Validation and equivalence 2006. The National Institute for Public Health and the Environment (RIVM), The Netherlands.
http://www.rivm.nl/bibliotheek/rapporten/680708001.html.

PAPERS

COMPARISON OF REGRESSION METHODS IN EQUIVALENCE DEMONSTRATION. Reinhold Beier

NEW ISSUES RELATED TO THE PM10 AND PM2.5 STANDARDS: TOWARDS A ROBUST CONTROL OF UNCERTAINTY. A. van der Meulen.

EXPERIENCES WITH EQUIVALENCE TESTING IN GERMANY. Ulrich Pfeffer, Thorsten Zang, Reinhold Beier and Eva-Maria Doutnik.

EQUIVALENCE OF PM10-MEASUREMENTS IN THE DUTCH NATIONAL AIR QUALITY MONITORING NETWORK Ronald Hoogerbrugge, Ruben Beijk.

PRACTICAL TESTING OF EQUIVALENCE DEMONSTRATION FOR PM AUTOMATED MONITORING METHODS: THE EXPERIENCE IN FRANCE. O. Le Bihan, F. Mathé, H. Marfaing, D. Robin

EQUIVALENCE OF DIFFUSIVE SAMPLERS TO REFERENCE METHODS: THE OZONE RADIAL DIFFUSIVE SAMPLER AND MEMBRANE-CLOSED PALMES TUBE EXAMPLES. Michel Gerboles, Daniela Buzica, Hervé Plaisance, A. Piechocki, G. Galloo

PM10 EQUIVALENCE TESTS IN SWITZERLAND: RELATIONSHIP BETWEEN A TEOM-FDMS AND THE REFERENCE METHOD AT A SUBURBAN SITE. Ch. Hueglin, R. Gehrig, Beat Schwarzenbach.

NON SEASONAL VARIABILITY AND THE EFFECT OF LOGGING AVERAGES ON TEOM AND FDMS MEASURMENTS MADE DURING THE UNITED KINGDOM PM_{10} and $PM_{2.5}$ EQUIVALENCE TRIALS. David Harrison, Jeff Booker and Richard Maggs.

COMPARISON OF REGRESSION METHODS IN EQUIVALENCE DEMONSTRATION

Reinhold Beier
Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen (Germany)
reinhold.beier@lanuv.nrw.de, Tel +4920179951474, Fax +4920179951575

Abstract

Regression lines obtained in evaluation of intercomparison exercises of PM10-measuring methods are validated by comparison with alternatives. A "posteriori" validation is based on robust regression diagnostics. Figures of merit known from Ordinary Least Square Regression like residual standard deviation and the coefficient of determination turn out inappropriate within this context. Instead, three robust statistics are used to assess properties of the distribution of residuals such as randomness, symmetry and coverage of the 95%-uncertainty range. The compared regression lines were provided by Ordinary Least Square Regression [1]**,** Orthogonal Regression [2] and Explorative Regression [3]. In an intercomparison exercise of automatic PM10-monitors with collocated High Volume Samplers, Orthogonal Regression and Explorative Regression were found preferable each in 37 % of the evaluated cases. Ordinary Least Square Regression turned out advantageous in the other 27 %.

Introduction

In intercomparison exercises for air quality measuring methods with a reference method, deviations may occur in results $y(j)$ of the candidate method, as well as in reference values $x(j)$ to be related by a statistical model equation $y(j) = a + b \cdot x(j) + e(j)$. For sake of simplicity, repeatability standard deviations of both methods are treated as constants not depending on the result of measurement. In this case, Ordinary Least Square Regression (OLS) [1] is not a-priori the most promising regression technique. Basic regression techniques allowing for errors in both variables are Orthogonal Regression ORT [2] and Explorative Regression EXP [3]. The main question addressed in this paper is, which of the compared regression techniques may be considered the (relative) best one for evaluating a considered series of observations. The compared regression methods are briefly summarized in "Compared regression methods". In "Robust regression diagnostic", the chosen approach to robust regression diagnostics is presented. In "Appications", a few applications are presented.

Compared regression methods

The regression methods compared are Ordinary Least Square Regression OLS [1], Orthogonal Regression ORT [2] and Explorative Regression EXP [3]. In the context of method intercomparison, an important task of data treatment is to establish a statistical model equation of type (1) relating two series of observations $x(j)$ and $y(j)$ with $j = 1$ to N knowing that both series may be subject to unknown deviations caused by imperfect measurement.

$$y(j) = a + b \cdot x(j) + e(j) \qquad (1)$$

Here, $x(j)$ designates the reference value provided by a reference method and $y(j)$ a result of measurement obtained by a single realization of the candidate method. If the reference method is represented by two or more collocated measuring systems of the same type, $x(j)$ may be obtained as

arithmetic mean value of these reference instruments. Reference values $x(j)$ with j = 1 to N are requested not to exhibit a common (systematic) deviation. The residuals $e(j)$ are expected to form a sample of a random variable of mean value zero, i. e. $\sum_{j=1}^{N} e(j) = 0$. Applicable rules for estimating slope b and intercept a as well as for estimating the corresponding standard uncertainties $u(b)$ and $u(a)$ are summarized in **Table 1**.

Orthogonal Regression ORT is realized here by slope estimate $b_{ORT} = (b_{OLS} + \frac{1}{B_{OLS}})/2$ with b_{OLS} being the estimate for model equation (1) provided by OLS-regression and B_{OLS} designating the estimate for the inverse model equation $x(j) = A + B_{OLS} \cdot y(j) + E(j)$ obtained by OLS-regression. As the slope of the wanted regression line is expected to be close to one, b_{ORT} provides a good estimate of the orthogonal regression line. The standard uncertainty $u(b)$ of the orthogonal slope estimate b_{ORT} is obtained here by Type-B estimate as proposed by [4]. Orthogonal Regression ORT is designed for residuals of Gaussian distribution not exhibiting outliers.

Explorative Regression EXP [3] is a basic robust regression technique neither requesting Gaussian distribution of residuals nor requesting absence of outliers. In Explorative Regression, slope estimate b is provided by an estimate of the median of possible slope-values $b(j) > 0$. The value $b(j)$ with j = 1 to N is determined by a straight line connecting data point $x(j)$, $y(j)$ with the centre of gravity exhibiting the coordinates $(\bar{x}, \bar{y})$.

Explorative slope b and the corresponding standard uncertainty $u(b)$ are estimated by robust statistics. For this purpose, the values $b(j) \le 0$ are discarded. The remaining $K \le N$ values $b(j) > 0$ are rearranged in increasing order and relabelled in such a way that $0 < b_{(1)} \le b_{(2)} \le .. \le b_{(M)} \le ... \le b_{(K)}$. Accordingly, $b_{(1)}$ designates the smallest positive value of the set $b(j)$ with j = 1 to N and $b_{(K)}$ the largest value. For odd numbers K, product $k = 0.5\ (K+1)$ is an integer and accordingly, slope estimate b is given by the sample value $b(j) = b_{(k)}$ with (integer) rank $k = INT(0.5 \cdot (K+1))$. For even numbers K, non-integer "rank" $k = 0.5 \cdot (K+1)$ shall indicate the well known interpolation rule for estimating the median by $b = [b_{(k\,\mathrm{int})} + b_{(k\,\mathrm{int}+1)}]/2$ with (integer) rank $k\,\mathrm{int} = INT(0.5 \cdot (K+1))$.

Robust regression diagnostics

Based on statistical model equation (1), the rules of uncertainty propagation [4, 5] provide relationship (2) for estimating the standard uncertainty $u(y(j))$ of results $y(j)$**.**

$$u(y(j)) = \sqrt{D^2 + s^2(e)(1 + \frac{1}{N}) + u^2(b)(x(j) - \bar{x})^2} \qquad (2)$$

$D = a + (b-1) \cdot x(j)$ designates the deviation of regression line $y'(j) = a + b \cdot x(j)$ about the ideal line $y'(j) = x(j)$. Deviation D is treated as a random effect of expected value zero exhibiting a fixed value throughout the evaluated intercomparison study. Expanded 95%-uncertainty $U_{95}(y(j))$ is provided by equation (3) with $k_{95} \cong 2.0$ for $N > 19$.

$$U_{95}(y(j)) = k_{95} \cdot u(y(j)) \qquad (3)$$

Table 1: Compared regression methods

y(j) = a + b x(j) + e(j)	***b; u(b)***	***a; u(a)***
Ordinary Least Square Regression (OLS)	$b = \frac{s(y)}{s(x)} \cdot R(x,y)$ $u(b) = \frac{s(e)}{s(x)} \sqrt{\frac{1}{N-1}}$	$a = \bar{y} - b \cdot \bar{x}$ $u(a) = \sqrt{\frac{s^2(e)}{N} + u^2(b) \cdot (\bar{x})^2}$
Orthogonal Regression (ORT) - approximation -	$b = \frac{s(y)}{s(x)} [R(x,y) + \frac{1}{R(x,y)}] / 2$ $u(b) = \frac{s(y)}{s(x)} \sqrt{\frac{1}{12} (R(x,y) - \frac{1}{R(x,y)})^2}$	$a = \bar{y} - b \cdot \bar{x}$ $u(a) \cong \sqrt{\frac{s^2(e)}{N} + u^2(b) \cdot (\bar{x})^2}$
Explorative Regression (EXP)	$b = median\{b(j) > 0\} = b_{(k)}$ $b(j) = \frac{y(j) - \bar{y}}{x(j) - \bar{x}}$ with j = 1 to N $u(b) = [b_{(m)} - b_{(n)}] / 2$ $0 < b_{(1)} \le b_{(2)} \le .... \le b_{(K)}$ $m = 0.5 \cdot (K + 1 + 1/\sqrt{K})$ $n = 0.5 \cdot (K + 1 - 1/\sqrt{K})$ K = number of $b(j) > 0$	$a = \bar{y} - b \cdot \bar{x}$ $u(a) \cong \sqrt{\frac{s^2(e)}{N} + u^2(b) \cdot (\bar{x})^2}$
Additional calculating rules		
$\bar{x} = \sum_{j=1}^{N} x(j) / N$ $\bar{y} = \sum_{j=1}^{N} y(j) / N$	$s(x) = \sqrt{\sum_{j=1}^{N} [x(j) - \bar{x}]^2 / [N-1]}$ $s(y) = \sqrt{\sum_{j=1}^{N} [y(j) - \bar{y}]^2 / [N-1]}$ $s(e) = \sqrt{\sum_{j=1}^{N} [y(j) - a - b \cdot x(j)]^2 / [N-2]}$	$R(x,y) = \frac{\sum_{j=1}^{N} [x(j) - \bar{x}][y(j) - \bar{y}]}{s(x) \cdot s(y)}$

In equation (2), a (constant) standard uncertainty $u(x)$ of the reference values $x(j)$ is taken into account implicitly. This is indicated by the relationship $s^2(e) = s_r^2(y) + b^2 \cdot u^2(x)$ with $s_r(y)$ describing repeatability standard deviation of the candidate method.

A regression line is considered preferable within a given set of regression lines, if it complies best with the following objectives:

a) The residuals $e(j)$ with j = 1 to N form a random sample.
b) The residuals $e(j)$ with j = 1 to N follow a symmetric distribution of expected value zero about the regression line.
c) The fraction of observed values $y(j)$ not covered by the 95%-range $[x(j) - U_{95}(y(j)); x(j) + U_{95}(y(j))]$ is 5%.

Appropriate robust performance characteristics are summarized in **Table 2**.

Table 2: Robust performance characteristics for regression lines.

Description	**Performance characteristic *PC***	**Expectation value *E(PC)***	**Standard deviation *s(PC)***
Measure of randomness PC_{ran}	$p(e(j)e(j-1)>0) =$ $portion\{e(j)\cdot e(j-1)>0\}$	0.50	$0.5/\sqrt{N}$
Measure of symmetry PC_{sym}	$p(e>0) =$ $portion\{e(j)>0\}$	0.50	$0.5/\sqrt{N}$
Fraction of data points not covered by 95%-range PC_{cov}	$p_{out} =$ $portion\{y(j)\le x(j)-U_{95}(y(j))$ $\_or\_\ y(j)>x(j)+U_{95}(y(j))\}$	0.05	$\sqrt{0.05\cdot 0.95/N}$

Performance characteristic $PC_{ran} = p(e(j)e(j-1)>0)$ is a robust measure of randomness of the sample of residuals *e*(*j*) with *j* = 1 to *N*. For quantifying the value of PC_{ran}, the data triples (*x*(*j*), *y*(*j*), *e*(*j*)) are labelled by index *j* in ascending order of *x*(*j*). The probability of two subsequent residuals *e*(*j-1*) and *e*(*j*) to exhibit the same sign is simply 50 percent in case of a random sample. Accordingly, the expectation value of random variable PC_{ran} is $E(PC_{ran}) = 0.50$ under the hypothesis that the set *e*(*j*) with *j* = 1 to *N* forms a random sample. The corresponding standard deviation is $s(PC_{ran}) = 0.5/\sqrt{N}$. Performance criterion $PC_{sym} = p(e>0)$ provides information on the degree of symmetry of the sample of residuals *e*(*j*) with *j* = 1 to *N*. The expectation value of random variable PC_{sym} is given by $E(PC_{sym}) = 0.50$ and the standard deviation by $s(PC_{sym}) = 0.5/\sqrt{N}$. Performance characteristic $PC_{cov} = p_{out}$ designates the fraction of data points (*x*(*j*), *y*(*j*)) not covered by 95%-uncertainty range $[x(j)-U_{95}(y(j)); x(j)+U_{95}(y(j))]$. The expectation value of random variable PC_{cov} is given by $E(PC_{cov}) = 0.05$ and the standard deviation by $s(PC_{cov}) = \sqrt{0.05\cdot 0.95/N}$.

For *N* > 20, the binomial distributions of the considered performance characteristics are sufficiently well approximated by Gaussian distributions. For each performance characteristic PC_i, a considered regression line is assigned a z-score z_i by means of relationship $z_i = |PC_i - E(PC_i)|/s(PC_i)$ resulting in a triple of z-scores z_1, z_2, z_3 characterizing a regression line. The maximum *max*(*z*) = *max*(z_1, z_2, z_3) is used as a figure of merit to characterize fitness for purpose of a considered regression line. In conclusion, the regression line exhibiting the smallest value *max*(*z*) is considered the (relative) best within the group of compared regression lines. If two or three regression lines are assigned the same value *max*(*z*), the line providing the smallest uncertainty estimate is preferred.

Figures of merit known from Ordinary Least Square Regression like residual standard deviation and the coefficient of determination turn out inappropriate for an objective assessment of different regression lines, because they always prefer Ordinary Least Square Regression.

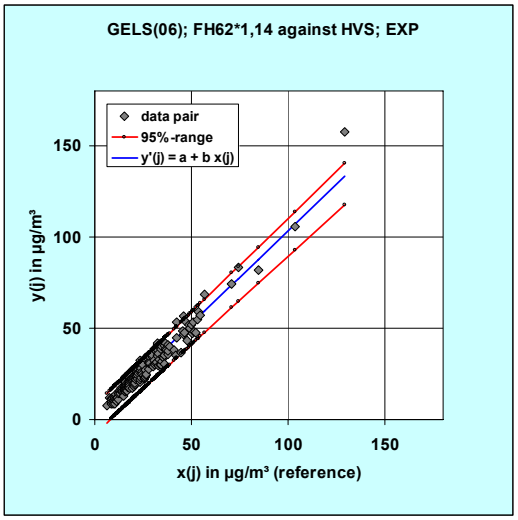
Applications

Series of observations were obtained in collocated application of automatic PM10-monitors of type beta-gauge (FH62-IR) and oscillating microbalance (TEOM) with High Volume Samplers (Digitel DHA-80 using quartz fibre filters) used as a local reference instruments. An intercomparison exercise of this kind was performed in 2006 inter alias for estimating correction factors for automatic PM10-monitors. An example of intercomparison of a beta-gauge PM10-monitor with a High Volume Sampler is summarized in **Table 3**. Daily mean values *y(j)* provided by beta-gauge instrument in Gelsenkirchen have been corrected by a factor of 1,14 before statistical evaluation.

With $max(z) = 0.7$, **Table 3** reveals Explorative Regression EXP as preferable compared with the regression methods OLS and ORT. Explorative Regression EXP scored best in all three performance characteristics. The resulting (relative) best estimate of 95%-uncertainty is $U_{95}(y_B) = 17.7$ % . In this example, 4 outlier were detected, but not rejected. Evidently, the presence of these outliers caused overestimation of the 95%-uncertainty by the regression techniques OLS and ORT. **Figures 1, 2** provide a graphical impression of differences between alternative regression lines.

Table 3: Intercomparison of corrected beta-gauge PM10-monitor operated in 2006 in Gelsenkirchen in collocation with a High Volume Sampler (Digitel DHA-80) without outlier rejection.

		Regression Method		
REGRESS_44	**Unit**	**OLS**	**EXP**	**ORT**
Site	–	GELS(06)	GELS(06)	GELS(06)
Pollutant	–	PM10	PM10	PM10
x	µg/m³	HVS	HVS	HVS
y	µg/m³	FH62*1,14	FH62*1,14	FH62*1,14
y_B	µg/m³	50	50	50
$u(y_B)$	µg/m³	5,0	4,5	5,4
$U_{95}(y_B)$	%	19,6	17,7	21,1
		$y(j) = a + b\,x(j) + e(j)$		
N		179	179	179
a	µg/m³	0,19	1,23	-0,48
b		1,06	1,02	1,09
$s(e)$	µg/m³	3,73	3,78	3,76
R(e,x)		0,00	0,16	-0,11
	Optimum	**Performance Characteristics**		
$p(e(j)e(j-1) > 0)$	0,50	0,54	0,51	0,56
$p(e(j) > 0)$	0,50	0,53	0,48	0,54
p_{out}	0,05	0,039	0,039	0,034
max(z)	0,0	1,1	0,7	1,6



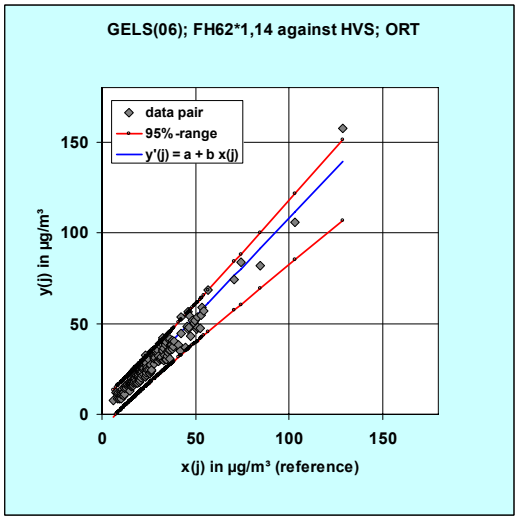


Figure 1: Explorative Regression line and 95%-range **Figure 2:** Orthogonal Regression line and 95%-range

In **Table 4,** the same series of observations was evaluated after rejection of four outliers. With $max(z)$ = 0.5 both, Explorative Regression EXP and Orthogonal Regression ORT were identified as equivalent. Due to the smaller 95%-uncertainty $U_{95}(y_B)$ = 14.3 % , Explorative Regression EXP is considered again the best regression method within the considered set. The decrease in the figure of merit FM from $max(z)$ = 0.7 to $max(z)$ = 0.5 after outlier rejection is an expression of the improvement in the statistical properties of the residuals achieved by outlier rejection.

Table 4: Intercomparison of corrected beta-gauge PM10-monitor operated in 2006 in Gelsenkirchen in collocation with a High Volume Sampler (Digitel DHA-80) after rejecting 4 outliers.

		Regression Method		
REGRESS_44	**Unit**	**OLS**	**EXP**	**ORT**
Site	–	GELS(06A)	GELS(06A)	GELS(06A)
Pollutant	–	PM10	PM10	PM10
x	µg/m³	HVS	HVS	HVS
y	µg/m³	FH62*1,14	FH62*1,14	FH62*1,14
y_B	µg/m³	50	50	50
$u(y_B)$	µg/m³	3,5	3,6	3,8
$U_{95}(y_B)$	%	13,8	14,3	14,9
		$y(j) = a + b\,x(j) + e(j)$		
N		175	175	175
a	µg/m³	1,74	1,43	1,13
b		1,00	1,01	1,02
$s(e)$	µg/m³	3,08	3,09	3,10
R(e,x)		0,00	-0,05	-0,11
	Optimum	**Performance Characteristics**		
$p(e(j)e(j-1) > 0)$	0,50	0,49	0,50	0,52
$p(e(j) > 0)$	0,50	0,49	0,52	0,51
p_{out}	0,05	0,069	0,057	0,057
max(z)	0,0	1,1	0,5	0,5

Discussion

A-posteriori validation of regression lines is an important issue (not only) in method intercomparison. The described procedure of robust regression diagnostic offers an attractive approach to objective a-posteriori assessment of deliberate regression lines. This procedure was evaluated in application to 30 data sets obtained in an intercomparison exercise of beta-gauge- and TEOM-PM10-monitors with High Volume Samplers (Digitel DHA-80) executed in 2006. In this evaluation, Ordinary Least Square Regression OLS was found preferable in 8 cases (27%), Explorative Regression EXP in 11 cases (37%), and Orthogonal Regression ORT in 11 cases (37%). A majority of evaluated data sets contained extreme data pairs (outlier) not following a Gaussian distribution. The described procedure of robust regression diagnostic may be applied as well for a-posteriori validation of any other regression line, e. g. obtained by Zero Intercept Regression, Weighted Orthogonal Regression or any other regression technique. A template (MS-EXCEL®) for application of the described regression techniques and the described procedure of regression diagnostics is available on demand.

References

[1] Draper, N. R. and H. Smith: *Applied Regression Analysis*, 3rd Edition, John Wiley & Sons, New York (1998)

[2] Isobe, T., Feigelson, E.D., Akritas, M.G., Babu, G.J.: Linear Regression in Astronomy I. The Astrophysical Journal, 364 (1990), 104–113

[3] Beier, R: Evaluation of method intercomparison tests by Explorative Regression, Gefahrstoffe – Reinhaltung der Luft, 67(2007), 49-54

[4] ISO-Guide 98, Guide to Expression of Uncertainty in Measurement, Geneva (1995)

[5] EN/ISO 20988, Air Quality – Guidelines for estimating measurement uncertainty, Geneva (2007)

NEW ISSUES RELATED TO THE PM_{10} AND $PM_{2.5}$ STANDARDS: TOWARDS A ROBUST CONTROL OF UNCERTAINTY.

A. van der Meulen.
National Institute for Public Health and the Environment RIVM, Bilthoven / Netherlands
ton.van.der.meulen@rivm.nl

Introduction

The current focus on equivalence almost "automatically" calls for:

- a closer look into the performance of the EU Standard Methods for PM (EN12341 for PM10 and EN14907 for PM2.5), and
- consideration of the possibility to introduce standards for Automatic Monitoring Systems AMS.

The issue of equivalence, i.e. whether a non-standard method gives results equivalent for regulatory purposes to EU Standard Methods, has been addressed by a separate EC working group for all pollutants in a comprehensive and uniform way.

For the evaluation of the "(lack of) comparability" between a non-standard and the standard measurement method (expressed as uncertainty) it is assumed that the relationship between the measurement results from both methods can be described by a linear relation, to be established by orthogonal regression. Subsequently, from this orthogonal regression the uncertainty from the non-standard method can be inferred.

As a first remark, obviously the "lack of comparability" between a non-standard and the Standard Method (SM) is determined *inter alia* by the performance of the SM. A lower SM uncertainty does improve the reliability of the equivalence claim.

Secondly, for most Automatic Monitoring Systems AMS a good comparability is observed at different specific sites and ambient conditions (R^2 values of 0.94 or higher).

Prompted by this observation, it could be worthwhile to consider standardization of AMS, including measurement and data treatment. For, it should be noted that for the greater part EU monitoring networks are employing Automated Monitoring Systems.

Current PM Standard Methods for PM10 and PM2.5: a comparison.

The main elements in the current PM Standard Methods are listed in the **Table 1**.

Table 1: Requirement of the PM Standard Methods

	PM10	**PM2.5**
	Related to CEN standard requirements	
Uncertainty	No	GUM
Equivalency	Envelope	EU Data Quality Objectives
Calibration Factor	No	Yes
STP vs Ambient	STP	Ambient
	Scientific issues	
Sampling Inlet	3 types *LVS* *HVS* *WRAC*	2 types *LVS* *HVS*
Filter	1 type *Quartz fibre*	4 types *Quartz fibre* *Glass fibre* *PTFE* *PTFE coated glass fibre*
Conditioning		
Weighing Room	Yes	Yes
On Site	No	Yes

Current PM Standard Methods for PM10 and PM2.5: Field experiences

Field experiences from the current PM SMs do show that their comparability is not self evident. It is thought that this is mainly due to variations in the SM's within their specifications (see also aforementioned **Table 1**):

- PM inlet system (more than 1 option)
- Filter type (different options)
- Conditioning of (blank and loaded) filter to standard relative humidity

Some examples.

PM inlet:

In the field evaluation program to support the standard method EN14907 (PM2.5) it was observed that the HVS implementation of the $PM_{2.5}$ standard method showed site dependent differences from -25 % to +14 % compared to the LVS standard method.

Filter type:

Different studies on different filter types showed mixed results, and even different brands of Quartz fibre filters were different.

Different Filter *types*:
- Teflon ≈ **(1.02 - 1.09)** × Quartz (VMM, BE)
- Cellulose Nitrate ≈ **(1.14 - 1.36)** × Quartz (VMM, BE)
- Cellulose Nitrate ≈ **(1.02)** × Quartz (STIMES-Wiesbaden, DE)

Quartz Fibre Filter *brand*:
- Pall Tissuquartz / Whatman QMA / S&S-QF20 (out of production):
  differences ranging: **+ 1.04 => + 1.18** (VMM, BE)
- Whatman QMA vs. S&S-QF20:
  no differences (RIVM-NL)

Conditioning to Relative Humidity:

When preconditioning *fresh blank filters* prior to sampling, it seems as if it takes several tens of days to equilibrate. This implies that equilibration of a fresh filter to 50 % RH has by far not been completed during a full sampling + weighing cycle. Consequently, a part (up till 100 - 200 µg) of the total weight increase of a loaded filter could be attributed to the fact that the conditioning is not yet completed, giving rise to possible effects up to several µg/m3.

As to *PM loaded filters*, various studies on RH hysteresis do show effects up to several hundreds of µg, i.e. several ug/m3 expressed as PM concentration

Revision of the PM standards

It should be self evident that

- Different implementations of Standard Methods (within their limitations)
- Following QA/QC prescriptions of the pertinent Standard Methods
- Different users starting from the same Standard Method

do result in Comparable/ Equivalent results.

But, apparently it is NOT, presumably due to (too) much freedom in the current PM standards, as shown before, notably regarding:

- PM inlet system (more than 1 option)
- Filter type (different options)
- Conditioning of (blank and loaded) filter to standard relative humidity.

This is an important conclusion, forcing to reconsider / revise the current PM standards.
Guiding principles in the current revision of PM standards could be:

- A robust uncertainty showing least dependence on the requirements of the standard
- without changing the PM measurand
  i.e.: *no discontinuity with present time series of concentration.*

Therefore, the following *way out* could be considered.

- Improve comparability by limiting the options:
  - Only 1 inlet: LVS
  - Only 1 Filter type: Quartz or Teflon or Teflon coated glass fibre

- Better scientific knowledge, and control of RH effects
  - Improve the pre-conditioning procedure of blank filters
  - Reconsideration of RH conditioning (stricter limits, or lower RH?)

Standardization of automated methods

As stated above: there is only one PM Standard Method, but there are several Automated Monitoring Systems in the EU monitoring networks.

Hence, this calls for a need for:

- Harmonization of **AMS-PM** measurements in EU
- Consistent Data Quality of such **AMS-PM** measurements

Starting from the previous arguments, it could be worthwhile to consider standardization of automated methods, including measurement and data treatment.

A possible approach would be the following procedure

- Type approval by **Manufacturer**
  - Laboratory test: is the AMS technically suitable?
  - Equivalence test only for common / specified ambient situations in EU (e.g. at especially designated super-site(s) ?)

- Equivalence test by **Member States MS**
  Is the AMS suitable for specific MS situation?

Advantages:

- Equivalence testing by manufacturers of their respective AMS at specified ambient conditions greatly facilitates the comparison of the AMS, not only under laboratory conditions but also under field ones.
- shared test burden between manufacturer & MS, so that equivalence could be determined for individual Member States with a much reduced set of trials.

A separate part of the standard could be devoted to harmonization of data treatment, e.g. the treatment of negative values, outliers, averaging, measurements below the detection limit etc.

Example: in cross border studies between the Netherlands and Germany or Belgium, up till a few µg/m3 differences in yearly averages of some 20 – 30 µg/m3 have been observed, as a result of differences in data treatment.

References

[1] PM10 – EN12341 (1998). Air Quality – Determination of the PM_{10} fraction of particulate matter – Reference method and field test procedure to demonstrate reference equivalence of measurement methods

[2] PM2.5 - EN14907 (2005). Ambient air quality – Reference gravimetric measurement method for the determination of the $PM_{2,5}$ - mass fraction of suspended particulate matter

[3] Report: EU 22341 EN (2006) "Field test experiments to validate the CEN standard measurement method for $PM_{2.5}$.
" (http://ies.jrc.cec.eu.int/fileadmin/H04/Air_Quality/EUR_22341_CEN_PM_2_5.pdf).

[4] Documents N 307-309 from CEN/TC264/WG15 on filter material (Vercauteren & van der Meulen).

[5] EN 12341 and 14907; Issues in weighing of PM blank filters (inter alia conditioning of fresh blank filters). J.A. Visser; GGD - Amsterdam / NL

[6] Document N304 from CEN/TC264/WG15 on RH conditioning of loaded filters (Quincey & Hafkenscheid).

[7] RIVM report 680500003/2006. A comparison of automated measurements of air quality near Valthermond (NL) in 2004 between the Netherlands and Lower Saxony (Germany). A. van der Meulen, P. Wolfs, D. de Jonge, M. Koester , E. Helmholz, H. Rienecker

[8] CompAir: Field Inter-Comparison of Air Quality Measurements between The Netherlands and North Rhine-Westphalia (Germany). Dave de Jonge, Anton van der Meulen, Ulrich Pfeffer , Reinhold Beier, Jürgen Friesel

[9] Simultaneous air quality measurements by VMM (Vlaamse Milieumaatschappij), Belgium and RIVM (Rijksinstituut voor Volksgezondheid en Milieu), The Netherlands: A comparison of the results of automated measurements of SO_2, NO, NO_2, O_3, CO and PM_{10} in Biest-Houtakker (NL) in the summer of 2005. E. Roekens, C. Matheeussen, N. Wintjens, R. Bogaert, R. Hoogerbrugge, A. van der Meulen

EXPERIENCES WITH EQUIVALENCE TESTING IN GERMANY

Ulrich Pfeffer, Thorsten Zang, Reinhold Beier and Eva-Maria Doutnik.
Landesamt für Natur, Umwelt und Verbraucherschutz (LANUV) Nordrhein-Westfalen, Recklinghausen (Germany)
ulrich.pfeffer@lanuv.nrw.de, Tel. +49-201-7995 1264, Fax +49-201-7995 1575

Introduction

Extensive equivalence trials have been conducted in Germany. These experiments were mainly focussed on PM measurement methods, but also diffusive samplers for nitrogen dioxide (NO_2) were studied.

Monitoring of PM

Experience across Europe shows that most 'classical' continuous PM10 monitoring methods underestimate real concentrations by approximately 10 to 30 percent. Therefore, data gained by these methods have to be corrected by a factor in order to meet the data quality objectives given by the first European Air Quality Daughter Directive (expanded 95%-uncertainty of 25 % for 24 hour values).

In the ambient air monitoring network LUQS in North Rhine-Westphalia, PM10 measurements are performed at about 70 stations. At 30 stations gravimetric High Volume Samplers (Digitel DHA-80; quartz fibre filters (QFF)) were operated in parallel to continuous monitoring methods (TEOM SES and β-gauge FH 62 IR with temperature control system TCS) in 2005. **Figure 1** shows the PM10 measurement stations of the LUQS monitoring network.

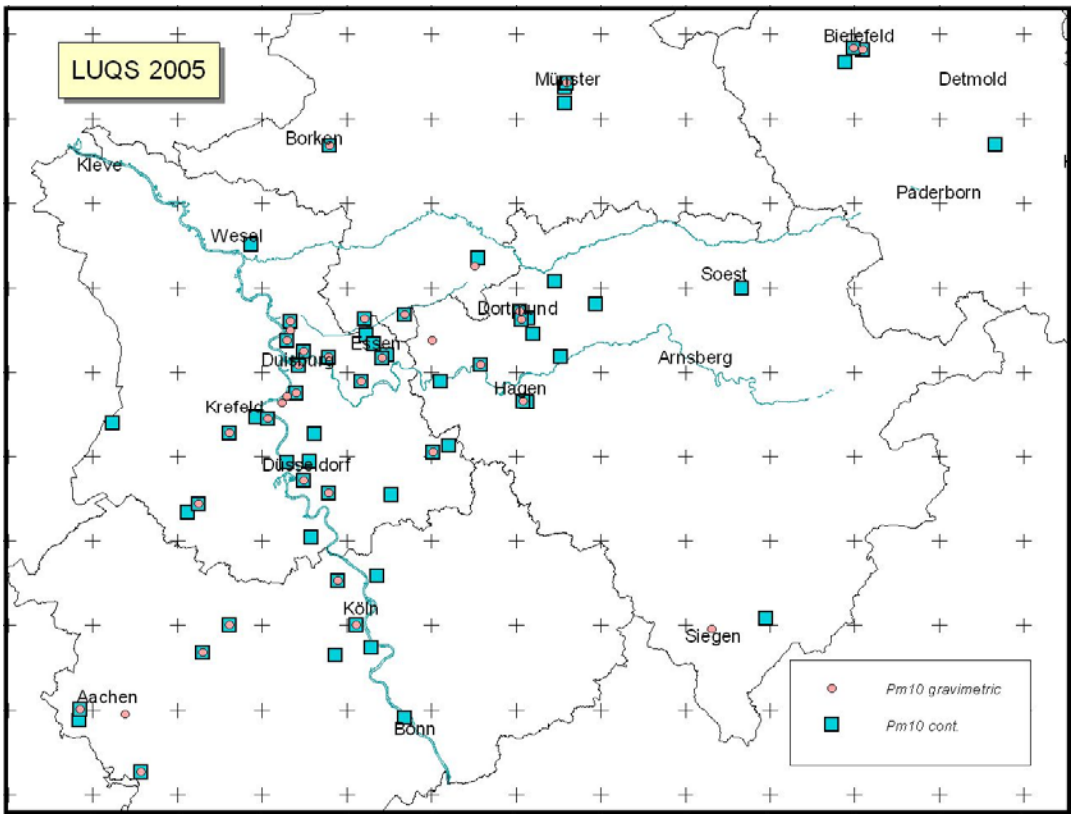


Figure 1: LUQS monitoring stations for PM10 in North Rhine-Westphalia (2005)

The evaluation of data gathered in previous years clearly shows that the gravimetric High Volume method is equivalent to the reference method described in EN 12341 (Low Volume Sampler). The expanded 95%-uncertainty at the limit value of 50 µg/m³ is lower than 12 percent. These results are documented in reports of the German ambient air monitoring networks [1] [2] [3]. The results of the Wiesbaden campaign [2] [3] were re-evaluated using the methods described in the European Guidance Document 'Demonstration of Equivalence of Ambient Air Monitoring Methods' (see **Figure 2**).

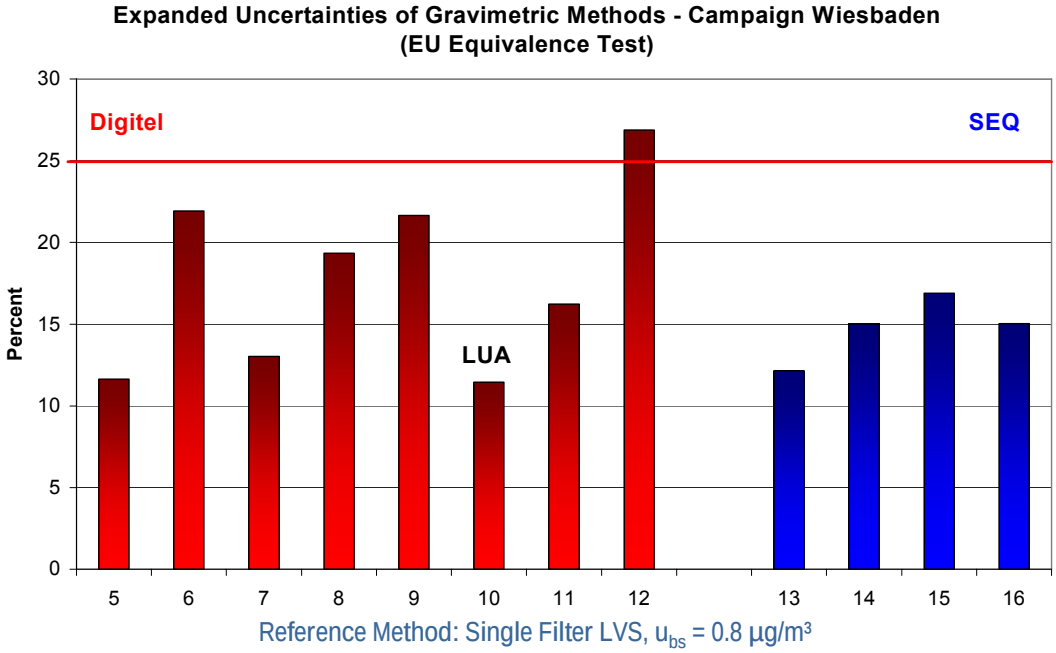


Figure 2: Expanded 95%-uncertainties of gravimetric PM10 monitoring methods in German monitoring networks

In a simple and pragmatic approach, general calibration factors of 1.26 ± 0.06 for TEOM SES and 1.14 ± 0.03 for FH 62 IR with TCS were calculated (2005). These factors apply for the majority of stations. In some cases site-specific corrections were necessary. For these corrected PM10 data expanded 95%-uncertainties were evaluated using the procedure laid down in the European Equivalence Report. For 90 percent of all stations with parallel measurements the data quality objective of 25 % is met after correction. The averaged expanded 95%-uncertainty of the corrected FH 62 IR data is 14.3 % and lower than that for the TEOM SES (22.2 %, including cases exceeding the data quality objective). The results for all stations are shown in **Figure 3** [4].

Data analysis shows that simple correction factors for continuous PM monitoring methods can be easily derived from parallel measurements of continuous methods with a gravimetric reference method and work very well for annual averages, but are not ideal for the correction of daily averages. However, it is explicitly stressed that in spite of these results correction factors are indispensable and shall be applied anyway.

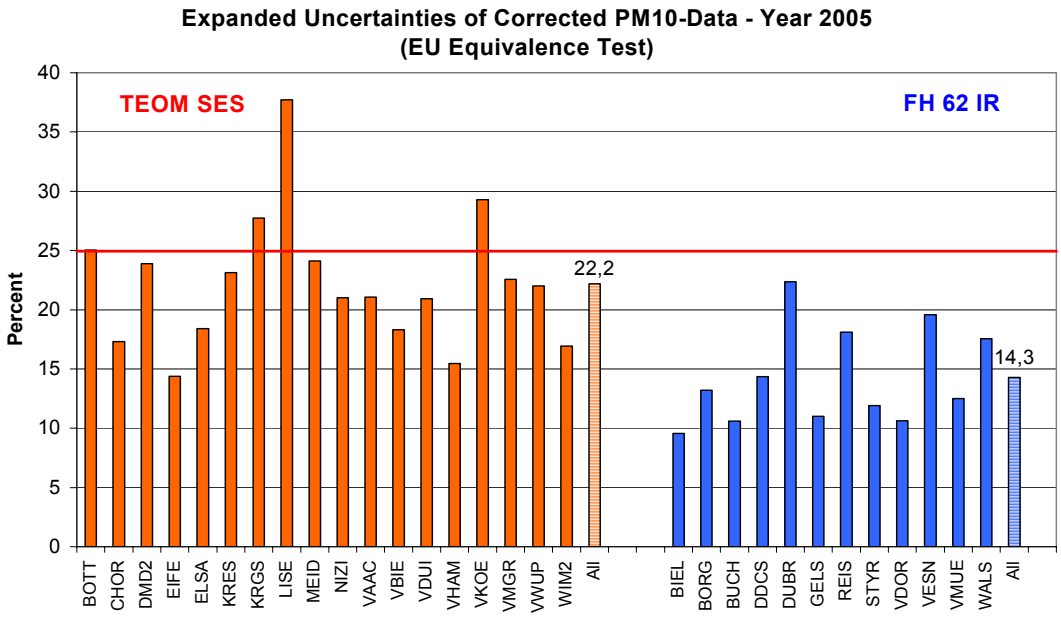


Figure 3: Expanded 95%-uncertainties of continuously measured PM10 data after application of calibration factors

Against the background of these results new developments in the field of continuous PM monitoring methods have been watched carefully. In North Rhine-Westphalia, the TEOM FDMS (Filter Dynamics Measurement System) was tested by LANUV (formerly: LUA) for PM10 following the requirements of the EU-guideline mentioned above. At a traffic-related site and an industry-related site in Duisburg two TEOM FDMS (candidate method, CM) were operated over more than 6 months in parallel to two Digitel High Volume Samplers as the reference method RM (202 and 173 daily data pairs). The results of this equivalence tests are very promising:

Between-samplers uncertainties were calculated for the gravimetric method and the TEOM FDMS, respectively. The requirements of the guideline for this parameter (2 µg/m³ and 3 µg/m³) were safely met. No outliers were removed from the whole data sets.

By orthogonal regression the following functions were evaluated:

Traffic site (VDUI): $CM = 0.96 * RM - 0.51$

Industry site (DUBR): $CM = 1.00 * RM - 2.54$

The expanded uncertainty on a 95 % level of confidence, calculated according the EU guidance document without any correction, is 13.9 % (VDUI) and 17.1 % respectively (DUBR) at the limit value (50 µg/m³) and therefore well below the data quality objective of 25 %. All details are summarized in **Table 1** [4].

Table 1: Results of equivalence tests of TEOM FDMS in North Rhine-Westphalia

Parameter	**Duisburg - VDUI**	**Duisburg - DUBR**
Type	Traffic	Industry
Period	22.07.05 – 28.02.06	14.07.06 – 31.01.07
Reference Method	Digitel HVS / QFF	Digitel HVS / QFF
Nr of Data Pairs	202	173
Mean of RM Data (µg/m³)	35,5	33,3
Range of RM Data (µg/m³)	10,9 – 146,6	11,6 – 88,3
Between-RM Uncertainty (µg/m³)	1,49	1,86
Between-CM Uncertainty (µg/m³)	1,59	1,10
Relation RM – CM	CM = 0,96 * RM – 0,51	CM = 1,00 * RM – 2,54
Expanded Uncertainty CM	13,9%	17,1%
After Correction (S-I)	11,2%	14,8%

Additional campaigns for equivalence tests were performed by TÜV Rheinland (Cologne) for other types of instruments, for example the Thermo Model 5030 SHARP Monitor. This monitor was tested for PM_{10} and $PM_{2.5}$ against Leckel LVS3 or SEQ47/50 samplers at four sites according to the equivalence guideline. The requirements were met at all sites for PM_{10} and $PM_{2.5}$. Expanded 95%-uncertainties (without correction) for PM_{10} are in the range of 9 % to 14 %, for PM2.5 in the range of 20 % to 23 %. Details may be found in [5].

Monitoring of Nitrogen Dioxide

Parallel measurements with modified Palmes type diffusive samplers and continuous monitors representing the reference method according EN 14211 were performed at ten mostly traffic-related measurement sites in North Rhine-Westphalia over 14 months in 2004-2005.

Classical Palmes tubes without any barrier against air turbulences did not give satisfactory results. In order to eliminate such effects, the samplers were modified with a glass frit at the inlet. Details can be found in [6].

The NO_2-uptake rate for the modified Palmes tubes was evaluated to be 0.8270 ± 0.0085 cm³/min and did not show a dependence on exposure interval and concentrations within reasonable limits. The evaluation is shown in **Figure 4** where the sampled mass of NO_2 is plotted against the dose (product of average concentration and exposure time). The slope of this regression line represents the uptake rate.

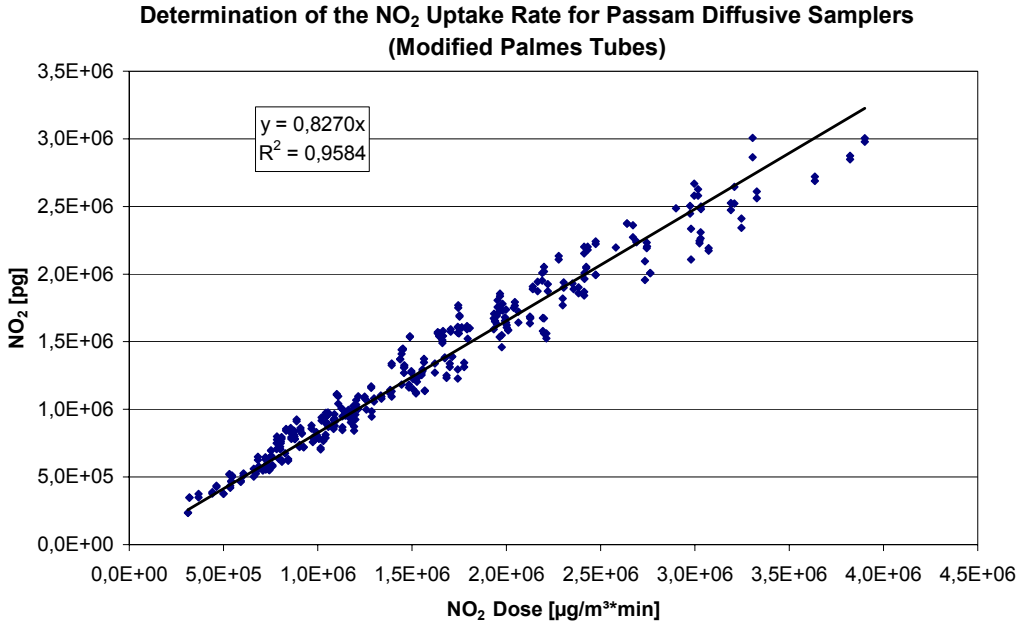


Figure 4: Evaluation of uptake rates for modified Palmes tubes

The data quality objective for measurement uncertainty of the first Daughter Directive of 25 % for indicative measurement is met for all evaluated exposure intervals between one and six weeks. For sampling periods of two weeks an expanded 95%-uncertainty of about 20 % at the EU-limit value of 40 µg/m³ can be achieved.

Results of an uncertainty evaluation based on the method described in the EU guideline report is summarised in **Table 2**.

Table 2: Results of EU-equivalence tests of modified Palmes tubes for NO_2 in North Rhine-Westphalia

REGRESSION OUTPUT (uncorrected data)		
slope b	*0,92*	significant
uncertainty of b	*0,02*	
intercept a	*4,76*	significant
uncertainty of a	*1,12*	
EQUIVALENCE TEST RESULTS		
random term	4,16	µg/m³
bias at LV	1,59	µg/m³
combined uncertainty	4,46	µg/m³
relative uncertainty at the LV	11,14	
RM between-sampler uncertainty	0,00	µg/m³

For this evaluation uncorrected data were used and no outlier test was performed. Because the reference method was not operated with two instruments in parallel no exact data for the between instrument uncertainty for this method are available. For this reason the between instrument

uncertainty was set to zero. As a consequence, the expanded uncertainty of the diffusive samplers is overestimated with 22 %.

Because the calibration of the diffusive samplers is based on the reference method of EN 14211 these results are principally unbiased. Since the comparisons were done at 10 different sites over 14 months most influences contributing to the uncertainty of the reference method including calibration are randomised and are contained in the evaluation of the average uptake rate of the modified samplers. For these reasons it seems to be justified to divide the uncertainty of single values by the square root of 12 when calculating annual averages from 12 monthly values. So the uncertainty of annual averages based on monthly measurements should be less than 10%. This means that the EU data quality objective of 15% is safely met for annual means.

The assumptions described above are verified by a comparison of annual averages measured by the reference method and diffusive samplers in 2006. It should be born in mind that this monitoring period in 2006 is totally independent of the period used for the evaluation of the uptake rate. In addition several new stations are included in this evaluation.

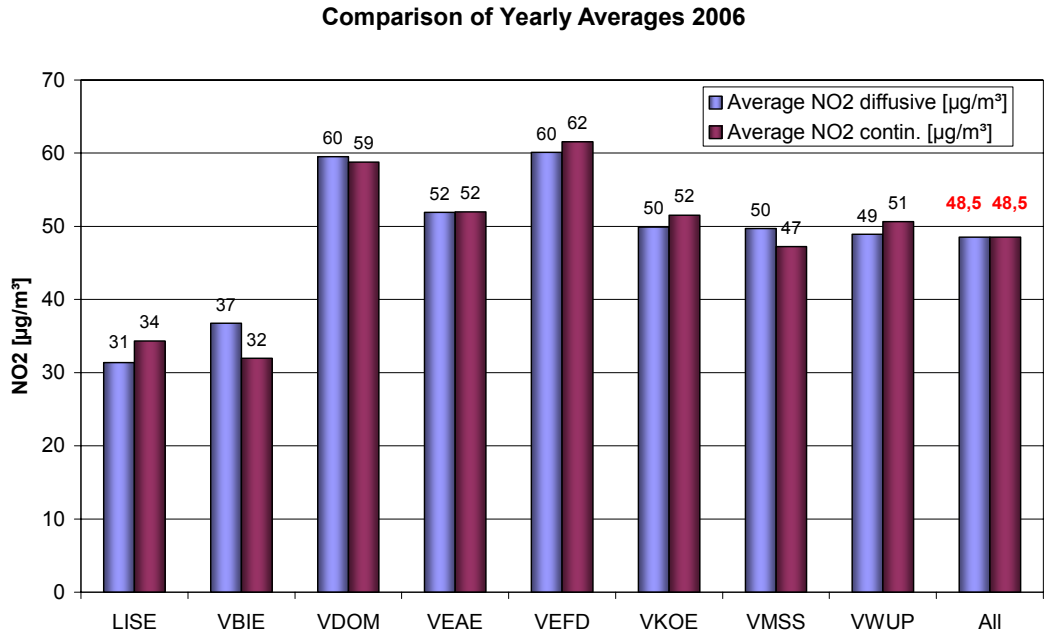


Figure 5: Comparison of yearly averages of NO_2 measured continuously and with diffusive samplers

Because of these results, Palmes tubes modified for the elimination of turbulence effects are an extremely effective and useful tool within the implementation of the first EU Daughter Directive. Hence, in North Rhine-Westphalia annual averages measured with the modified Palmes tubes are used for compliance monitoring and air quality plans.

Conclusion

Various equivalence trials were performed in the LUQS monitoring network in North Rhine-Westphalia (Germany). Continuous PM10 measurement with the β-gauge FH 62 IR and the TEOM SES were calibrated with the reference method represented by high volume samplers Digitel DHA-80. After correction the data quality objectives of the first daughter directive of 25% are met in 90% of all cases.

For the new instrument TEOM FDMS expanded 95%-uncertainties in the range of 14% to 17% are found without any data correction.

Annual means of nitrogen dioxide measured with modified Palmes tubes are shown to meet the data quality objective of 15% for the expanded 95%-uncertainty. Therefore, this method is applied also for compliance monitoring.

References

[1] Bericht zum Erfahrungsaustausch über "Messungen von PM10/PM2,5" im Auftrag des LAI Hessische Landesanstalt für Umwelt (HLfU), Wiesbaden, Landesumweltamt Nordrhein-Westfalen (LUA NRW), Essen, Gesellschaft für Umweltmessungen und Umwelterhebungen (UMEG), Karlsruhe, Karlsruhe (2000)

[2] H.-U. Pfeffer, R. Beier, J. Geiger, G. Löschau, W. Travnicek:. PM10-Vergleichsmessungen der deutschen Bundesländer mit gravimetrischen und kontinuierlichen Verfahren KRdL-Experten-Forum Staub und Staubinhaltsstoffe, 10./11. November 2004, KRdL-Schriftenreihe 33, S. 171-186 (2004)

[3] R. Beier, F. Bunzel, E. Klasmeier, L. Laskus, G. Löschau, M. Lohberger, R. Lumpp, F. Palme, U. Pfeffer, W. Travnicek:. PM10-Vergleichsmessungen der deutschen Bundesländer im Rahmen der STIMES-Arbeitsgruppe PM10 Materialien Band 66, Ed.: Landesumweltamt NRW (LUA NRW), Hessisches Landesamt für Umwelt und Geologie (HLUG), Umweltbundesamt (UBA), Landesumweltamt Essen 2005, ISSN 0947-5206 (Materialien), 175 pages

[4] U. Pfeffer: Application of Correction Factors for PM10-Measurements in North Rhine-Westphalia (Germany). Conference 'Particles in Europe', June 13-14, 2006 in Antwerp (Belgium)

[5] TÜV Rheinland Immissionsschutz und Energiesysteme GmbH, TÜV-Berichte 36/21203481/A and 936/21203481/B, Köln, 6.12.2006

[6] U. Pfeffer, R. Beier, T. Zang: Measurements of Nitrogen Dioxide with Diffusive Samplers at Traffic-Related Sites in North Rhine-Westphalia (Germany) - Performance Characteristics and Measurement Uncertainty. Gefahrstoffe – Reinhaltung der Luft 66 (2006), Nr. 1/2, 38-44

EQUIVALENCE OF PM_{10}-MEASUREMENTS IN THE DUTCH NATIONAL AIR QUALITY MONITORING NETWORK

Ronald Hoogerbrugge, Ruben Beijk
Laboratory for Environmental Monitoring, National Institute for Public Health and the Environment, The Netherlands
Email: Ronald.Hoogerbrugge@rivm.nl, Phone 0031-(0)302743306

Abstract

In the Dutch National Air Quality Monitoring Network (NAQMN) particulate matter (PM_{10}) is measured at various locations across the Netherlands. An equivalence study has been carried out in order to ensure the quality of these PM_{10}-measurements

As part of the equivalence study, automatic PM_{10}-measurements and the EU reference method are compared. The automatic PM_{10}-measurements in the Dutch National Air Monitoring Network have been calibrated based on the outcome of this study. The demonstration of equivalence between the automatic and reference method is performed conform the recommendations of the Clean Air For Europe (CAFE) steering group.

In order to obtain a consistent dataset the historic measurement results were recalculated according to the appropriate calibration functions. Also the historic data needed a transformation from the previously used standard conditions to the currently prescribed ambient conditions.

For the equivalency calculations orthogonal regression was used. For situations with an insignificant intercept orthogonal regression without intercept is applied. The equations for orthogonal regression without intercept and corresponding uncertainty were not available in the equivalence guideline. Such an approach was considered necessary therefore the statistical equations for orthogonal regression without intercept are presented and applied.

Methods

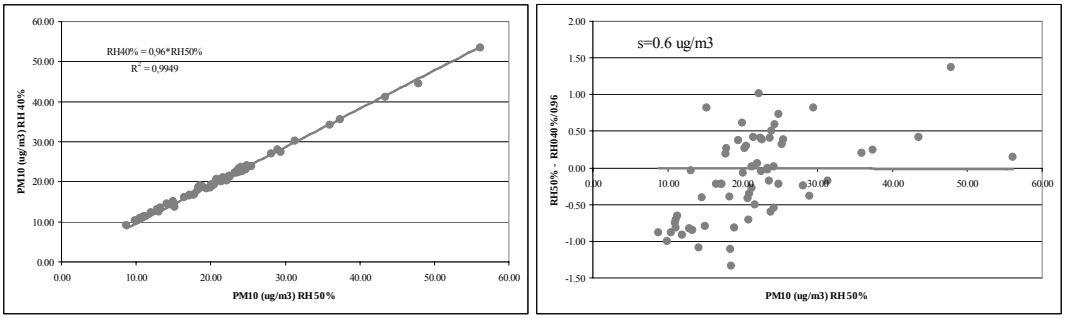
Reference measurements

Simultaneously with the automatic measurements a Low Volume Sampler with quartz filters is used to determine the reference values, following the EN12341 document. The standard uncertainty of the reference measurements is estimated to be 1.6 $\mu g/m^3$.
The relative humidity (RH) in the weighing room in the period prior to April 2006 appeared to be approximately 43 %. An experiment has been carried out to determine the consequences of this deviation from a RH of 50 %. The result of this experiment shows an average loss of 4.4 % between measurements carried out at 40 % instead of 50 % RH with an R^2 of nearly one as shown in **Figure 1**. Assuming a linear dependence on RH, a correction of 3 % (0.7 * 4.4 %) is required to compensate the loss of weight during gravimetric measurements carried out at 43 % instead of 50 % RH. The correction has a contribution to the standard uncertainty of approximately 0.6 $\mu g/m^3$.

PM_{10} measurements in the NAQMN

In the Dutch National Air Quality Monitoring Network (NAQMN) the number of automated PM_{10} monitoring sites expanded from approximately 20 to 40 in the period 2003 up to 2006. In 2006, roughly half of the monitoring sites contained an old monitor model (FH62 I-N), the other half the new model (FH62 I-R). The old models in the monitoring network will be replaced with the new type in 2007 and 2008.

Distinction is made between both regional and urban sites as well as the old and new monitor model (respectively FH62 I-N and FH62 I-R). The former is necessary due to possible difference of particle composition, the latter is due to differences in the preliminary treatment of the airflow. Because the old model is operated using two different heating configurations, distinction is made between the FH62 I-N with old (static 50 degrees Celsius) and new (atmospheric temperature plus 10 degrees Celsius) heating too. Especially with the old monitor type problems with operational settings were found [1].



a **b**

Figure 1: Comparison of weighing results with a RH of 50 % (x-axis) with the weighing result of the same sample at a RH of 40 %. Figure 1b shows the deviations from the regression line indicating a contribution of 0.6 $\mu g/m^3$ to the standard uncertainty.

Calibration and equivalence of automatic PM_{10}-measurements

In total, 763 day average samples at 8 different locations are included for the comparison between the reference method and the new (automatic) model at regional locations. For urban sites the number of samples is 463, measured at 4 different locations. Reference measurements for comparison with the old monitor model are only available at urban sites, in total 181 samples at 2 locations for those with original static heating and 239 samples at 2 locations for the old model with a new heating configuration.

Equivalence calculations

The equivalence study is based upon the recommendations of the Clean Air For Europe (CAFE) steering group [2]. Hence orthogonal regression is applied to determine the relation between the automatic and reference measurements for concurrent monitor configurations.
The equivalence document [2] describes in Annex B the equations to determine the orthogonal regression parameters a and b and their uncertainty. For the equivalence data of some monitor types the intercept was not significant. Then orthogonal regression without intercept is applied.
The equivalence document enables a rather inconsistent approach of orthogonal regression without intercept. In this approach the slope is calculated from the model with intercept. This inconsistency is recognized by the document and it's authors. However at that time no alternative was present to

calculate of orthogonal regression without intercept including an estimate for the uncertainty in regression slope.

In the astronomical literature some general equations were presented by the Isobe *et al.* [3] which could be transformed to the orthogonal regression without intercept and its associated uncertainty. The equations are derived in general in [3] in appendix A. From these equations the orthogonal regression without intercept is a special situation were all sums of squares are calculated without the subtraction of the origin. For orthogonal regression without intercept (a=0) the slope b can be calculated using an equation similar to Annex B of [2]. e.g. $S_{xx} = \sum (x - \overline{x})^2 \rightarrow S_{xx} = \sum x^2$ etc:

$$b = \frac{S_{yy} - S_{xx} + \sqrt{(S_{yy} - S_{xx})^2 + 4S_{xy}^2}}{2S_{xy}} \tag{1}$$

with

$$S_{xx} = \sum x_i^2, S_{yy} = \sum y_i^2, S_{xy} = \sum x_i y_i \tag{2}$$

The variance of the slope is estimated as:

$$u^2(b) = \frac{b^2}{4\frac{S_{xy}^2}{S_{xx}^2} + (\frac{S_{yy}^2}{S_{xx}^2} - 1)^2} (\frac{V_1}{S_{xy}^2} + 2\frac{V_{12}}{S_{xx}S_{xy}} + \frac{V_2}{S_{xx}^2}) \tag{3}$$

with:

$$V_1 = \sum x_i^2 (y_i - x_i (S_{xy} / S_{xx}))^2$$
$$V_{12} = \sum x_i y_i (y_i - x_i (S_{xy} / S_{xx}))(y_i - x_i (S_{yy} / S_{xy}))$$
$$V_2 = \sum y_i^2 (y_i - x_i (S_{yy} / S_{xy}))^2$$

The validity of equation 3 was confirmed by a bootstrap simulation. In this simulation alternative data sets are sampled from the original data set. The spread in the results of the alternative data sets equals equation (3). The uncertainty equation in Annex B of the equivalence document, with the adjusted sums of squares ($S_{xx} = \sum (x - \overline{x})^2 \rightarrow S_{xx} = \sum x^2$ etc.), yields comparable results with Equation 3. These results together with an analytical solution derived with the maximum likelihood method for the variance of a slope forced trough the intercept show all similar results [4].

After calibration the standard uncertainty is calculated using eq.(9.17) from [2] for orthogonal regression and eq.(9.15) using eq.(3) for the orthogonal regression without intercept. In both calculations the uncertainty of the difference in relative humidity (**Figure 1**) are added explicitly to the uncertainty budget. The standard uncertainty is then multiplied by 2 to arrive at the expanded uncertainty and divided by 50 µg/m^3 to calculate the relative uncertainty at the relevant limit value. The number and influence of outliers appeared to be negligible.

Results

The regression results are demonstrated in **Figure 2** and **3** for each monitor configuration separately. The results for the new model at regional sites shows a slope of 1.17 with a 2.7 µg/m^3 offset, and is the only category with a statistical significant intercept. The relative (measurement) uncertainty for this category is 17 %. The same model at urban sites shows a 1.20 slope based on orthogonal regression forced trough the origin, with a relative uncertainty of 17 %.

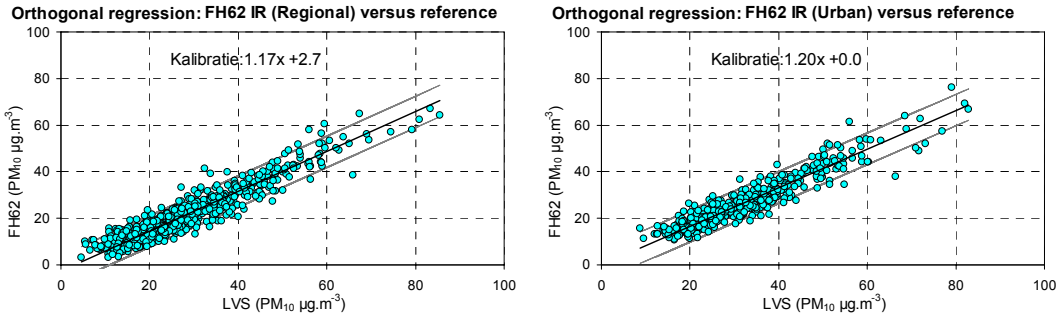


Figure 2: Equivalence data and calibration curves for the FAG62-IR monitor. The left hand figure shows the regional data and the right hand figure the data obtained on urban locations. In both figures the x axis shows the reference value and the y axis value from the automatic monitor.

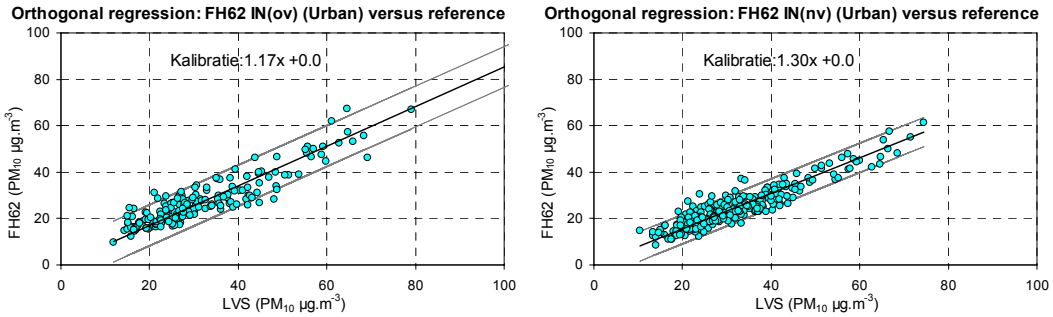


Figure 3: Equivalence data and calibration curves for the FAG62-IN monitor. Both figures show results of urban locations. The left hand figure shows the result of the original heating system the right hand figure the data obtained new heating system. In both figures the x axis shows the reference value and the y axis value from the automatic monitor.

The old monitor type (FAG-IN) at urban sites shows a slope of 1.17 and 1.30 for the configuration with the old and new heating treatment. The relative uncertainty of the former is 21 %, and 17 % for the latter. The European maximum allowed uncertainty is 25%, hence all four categories comply with this demand.

Definition of PM_{10} parameter

Before 2003, and in some cases up to 2005, particulate matter concentrations were reported using standard conditions (20 degrees Celsius and 1013 hPa). European legislation states PM_{10}-concentrations to be based on prevailing atmospheric conditions. Consequently, historic data is recalculated to meet this demand and to ensure consistency in trends. Measurements reported using standard conditions are therefore recalculated. Neither atmospheric temperature nor pressure were simultaneous available with all PM_{10}-measurements. Therefore meteorological data from the Dutch

Royal Meteorological Institute (KNMI) is used to recalculate the PM_{10}-measurements. Although the distance between KNMI-stations and PM_{10}-monitors may differ, the effect of this distance on the recalculated concentrations is negligible. The revalidation is based on the ideal gas law:

$$PM10_{nieuw}(t) = PM10_{oud}(t) \times \frac{p(t)}{1013} \times \frac{293}{273 + T(t)}$$

For each measured hour average *t*, PM_{10} ($\mu g/m^3$) concentrations are corrected with the environment and standard pressure (*hPa*) ratio, and, standard and environment temperature ratio (*degrees Celsius*).

New calibrations have been determined for both regional and urban situated sites based on the regression results, and replace the default factor of 1.3. The year average PM_{10}-concentrations before and after applying this new calibration is illustrated in **Figure 4**.

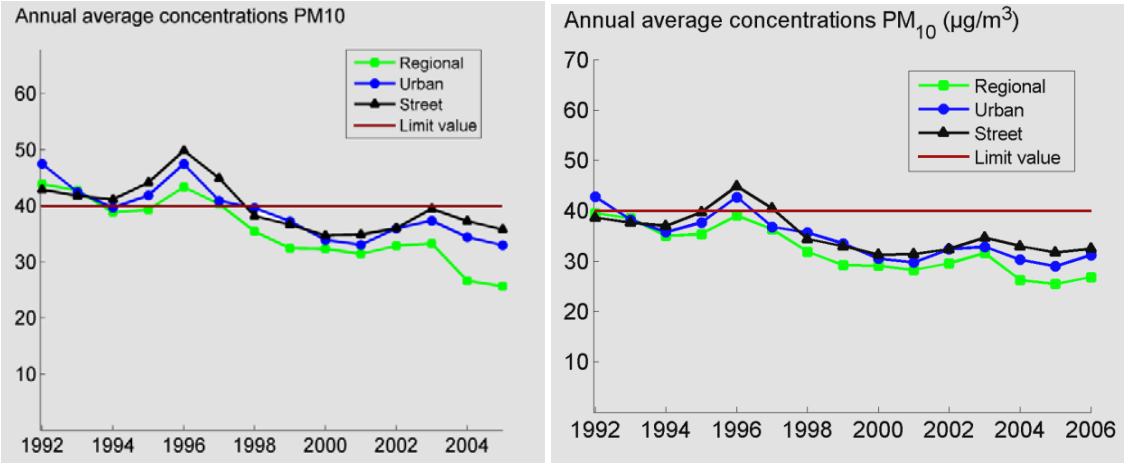


Figure 4: Annual average PM_{10}-concentrations based on former default calibration (left) and the new calibration as determined in the equivalence*/-*/- study (right).

Conclusion

Equivalence has been demonstrated for 4 configurations of one brand of β-attenuation monitor for PM_{10}. The results of the equivalence study reduce the uncertainties in PM_{10} measurements and make them comply with the EU required quality requirements. Historic PM_{10} measurements are recalculated to comply with the current definition of the parameter. For the equivalency calculations always orthogonal regression was used. For situations with an insignificant intercept orthogonal regression without intercept is applied. The equations for orthogonal regression without intercept and corresponding uncertainty were not available in the equivalence guideline. Such an approach was considered necessary therefore the statistical equations are presented and applied.

Acknowledgements

The vast amount of quality comparison data could only be obtained by the skilled effort of many colleagues of the National Air Quality Monitoring Network. Measurement phenomena were interpreted and solved by A. van der Meulen, F.Th. van Arkel, T.L. Hafkenscheid, G.C. Stefess and J.H. Verboom. For the statistical evaluation the assistance of D. Mooibroek and J. van de Kassteele is acknowledged.

References

[1] Beijk, R., Hoogerbrugge, R., Hafkenscheid, T.L., Arkel, F.Th. van, Stefess, G.C., Meulen, A. van der, Wesseling, J.P., Sauter, F.J., Albers, R.A.W. (2007) PM_{10}: Validatie en equivalentie 2006 (in Dutch). Bilthoven: National Institute for Public Health and the Environment (RIVM). Publication 680708001

[2] EC Working Group on Guidance for the Demonstration of Equivalence (2005) Demonstration of equivalence of ambient air monitoring methods. Published and available on the internet: http://ec.europa.eu/environment/air/ambient.htm (February 2007)

[3] T. Isobe, E.D. Feigelson, M.G. Akriteas and G.J. Babu, "Linear regression in Astronomy" Astroph. J. 364 104-113 (1990).

[4] Beijk, R. et al. "PM10: Equivalence study 2006" (2008). Bilthoven: National Institute for Public Health and the Environment (RIVM). Publication 680708002

PRACTICAL TESTING OF EQUIVALENCE DEMONSTRATION FOR PM AUTOMATED MONITORING METHODS: THE EXPERIENCE IN FRANCE

O. Le Bihan [1] , F. Mathé [2], H. Marfaing [3], D. Robin [4]
1) LCSQA-INERIS, Parc technologique Alata - 60550 Verneuil-en-Halatte, France (Olivier.Le-Bihan@ineris.fr)
2) LCSQA-Mines de Douai, 59508 Douai, France (mathe@ensm-douai.fr)
3) Airparif, 7 rue Crillon, 75004 Paris, France (helene.marfaing@airparif.asso.fr)
4) Airmaraix, 67-69 Avenue du Prado 13286 Marseille, France (dominique.robin@atmopaca.org)

Abstract

The present study summarizes the results of measurement campaigns conducted in France (Paris, Marseille), in Belgium (Aarschot) and in Italy (Monterotondo), in the framework of Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) activities. The main goal is to show that the methods tested (TEOM-FDMS from Thermo R&P and beta gauge MP101M-RST from Environnement SA) meet the Data Quality Objectives for PM10 continuous measurements specified in the Air Quality Directive 99/30/EC, under conditions reflecting practical application in air quality monitoring networks. $PM_{2.5}$ continuous measurement capability for TEOM-FDMS has also been studied.

Keywords: Demonstration of equivalence, Data Quality Objective, EU Directives, automatic monitors

Introduction

In France, measurement of the mass of particulate material in the atmosphere (mainly as PM10, but including some $PM_{2.5}$ monitoring) is carried out by 36 air quality monitoring networks with nearly 450 SPM automatic monitoring stations. These networks are operated at both the regional and local levels, on behalf of the French state, the local authorities, industries and environmental organisations. Since the 90's, automatic monitoring of particulate matter has been largely founded on the TEOM and beta attenuation analysers (85% vs 15%). The main reasons of these choices are metrological (widely used sensitive methods, data communication « on a daily basis » is possible…).

The present study summarizes the results of measurement campaigns conducted in the framework of LCSQA activities. The main objective is to show that the methods tested (TEOM – FDMS from Thermo R&P and beta gauge MP101M-RST from Environnement SA) meet the Data Quality Objectives for fixed measurements specified in the Air Quality Directive [1], under conditions reflecting practical application in air quality monitoring networks. The principles and methodologies to be used for checking the equivalence of non-reference methods for PM-monitoring are given in a document written by European Commission Working Group [2].

Methodology

Two experimental sites are located in France (Bobigny and Marseille) in an urban background area and near an air pollution monitoring station. The 2 other sites are located abroad (Belgium–Aarschot and Italy [3] Monterotondo) in suburbs (**Figure 1**):

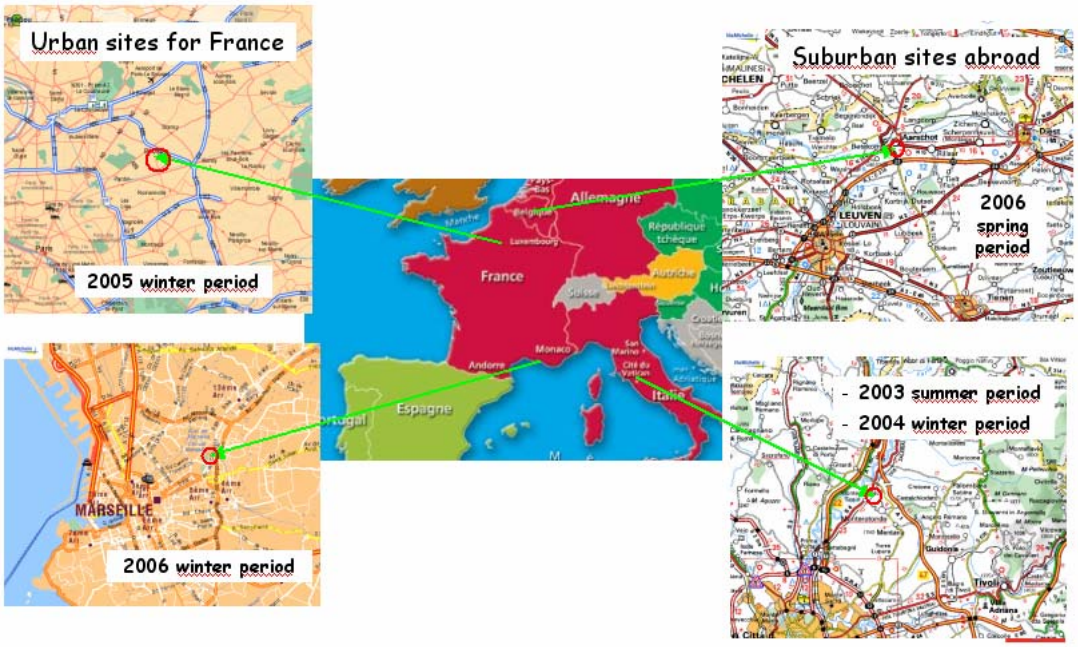


Figure 1: general overview of the 4 sampling sites situation

The choice of sites and time period of the year has been based on representativeness for typical conditions for which equivalence will be claimed, including possible episodes of high concentrations. These field tests shall be performed in which all methods are compared side-by-side (Fig.2). In each case, the experimental sites are located in an urban background or suburban area and near an air pollution monitoring station from an air quality network.

The choice of site and time period of the year has been based on representativeness for typical conditions for which equivalence will be claimed, including possible episodes of high concentrations.

Concerning particulate matter PM_{10} and $PM_{2.5}$, the gravimetric reference used for French campaigns was a Partisol Plus 2025 sequential sampler (Thermo R&P) collecting on 47 mm diameter, 2 µm pore size PTFE filters (Zefluor^{TM} from Pall Corporation).

Concerning fields tests performed in Belgium and Italy, only the beta attenuation monitor has been studied for the PM_{10} size fraction. The gravimetric reference used was respectively a Leckel SEQ 47/50 sequential sampler (Thermo R&P) collecting on 47 mm diameter, Quartz Fiber filters (from Macherey-Nagel) and a Reference Sampler designed by CNR-IIA collecting on 47 mm diameter, 2 µm pore size PTFE filters (TefloTM from Pall Corporation).

Concerning Bobigny and Marseille trials, the weighing procedure was conducted by INERIS according to requirements of European Standard EN 14907 [4], using a balance with a resolution of 10 µg in a temperature (20 ± 1°C) and humidity (50 ± 5% RH) controlled weighing room. For trials from abroad, the weighing operations were performed respectively by VMM and CNR-IIA. All apparatus were duplicated and equipped with same size-selective inlet as the reference method. Sampling and monitoring series have had a duration sufficient to collect a minimum of 40 duplicated pairs of measurement results each averaged over at least 24-hour per comparison. The main learning of these tests are the importance of apparatus configuration (step time measurement, data calculation, same inlets for all systems…) and the useful load of work in terms of QA/QC (filters temperature control throughout all measurement process, weighing procedure in T/RH controlled room, well documented operating procedures for cleaning of inlets, flow & leak control, data validation…)

Figure 2: General view of disposals at Bobigny (France)

Results

Figures 3 to **5** gives examples of results obtained during field tests (time series of PM_{10} concentrations, comparison of PM_{10} beta gauge and PM2.5 TEOM-FDMS with reference method):

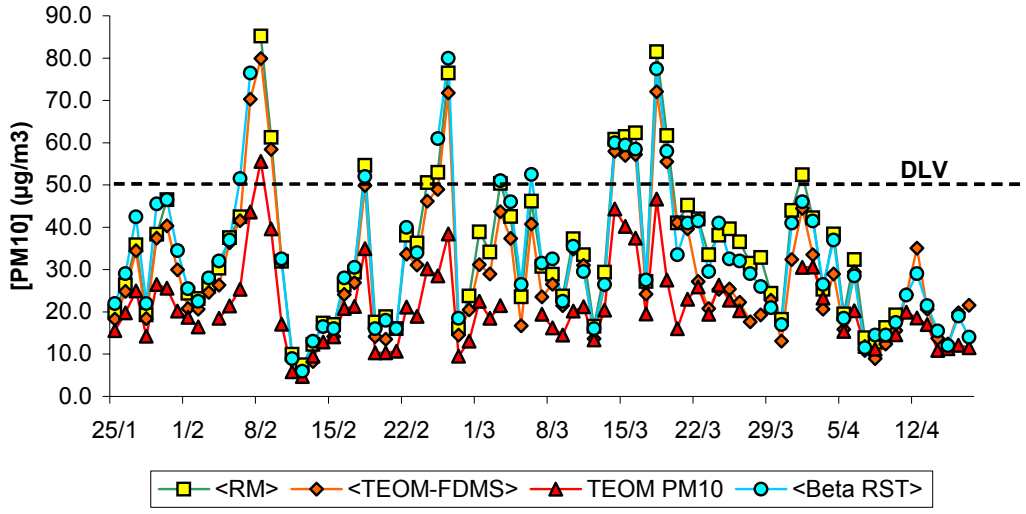


Figure 3: PM_{10} profiles at Bobigny

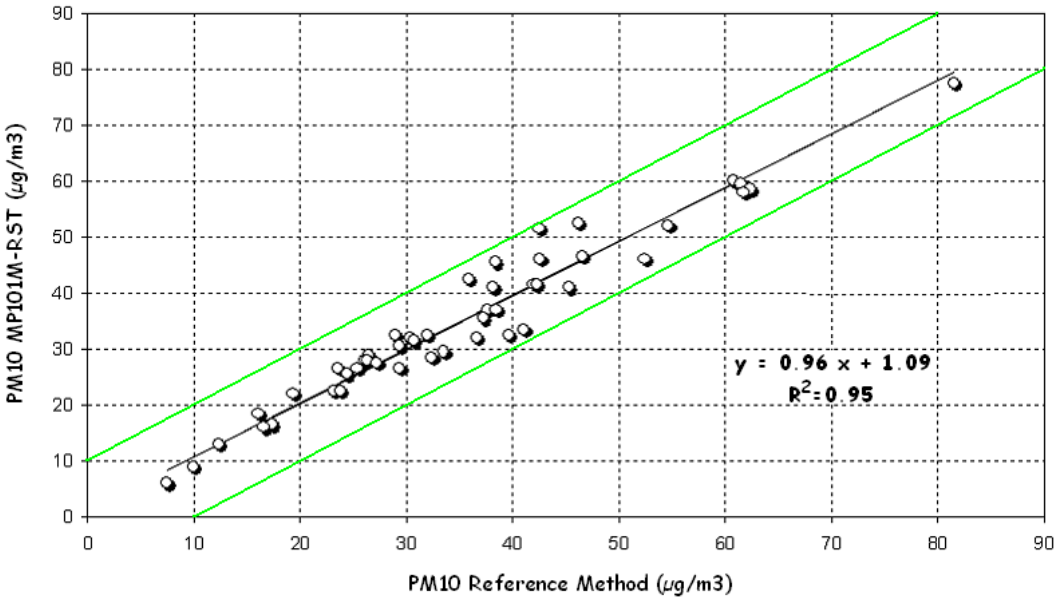


Figure 4: Comparison of β-attenuation monitor with reference method (PM_{10} 24h values)

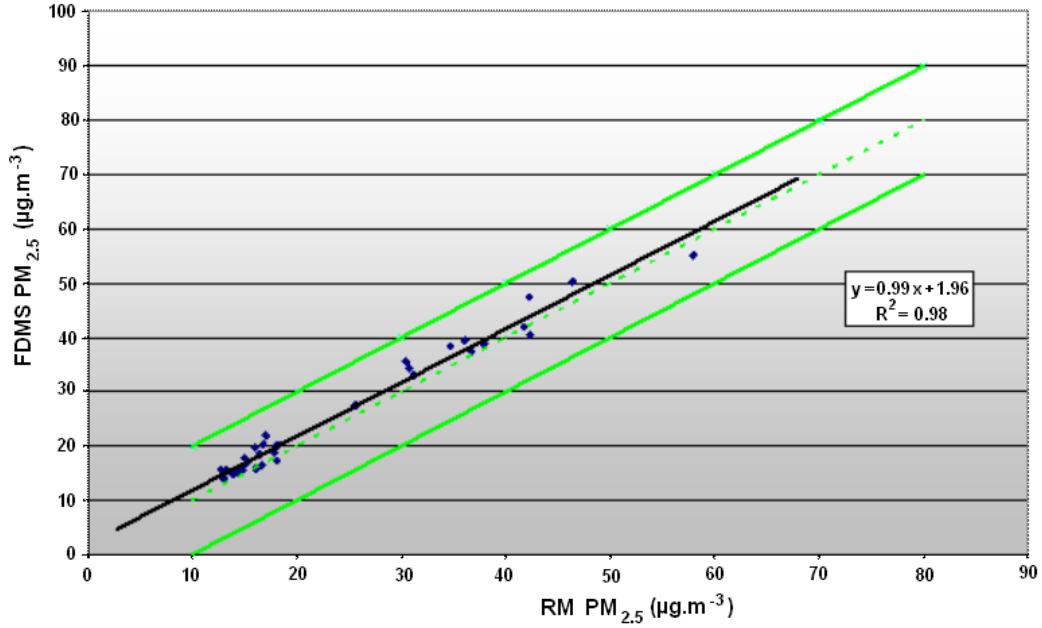


Figure 5: Comparison of TEOM-FDMS with reference method ($PM_{2.5}$ 24h values)

A good accordance of both new technologies with Reference Method is observed. two important points should be stressed:
- same exceedances number of daily limit value (50 $\mu g.m^{-3}$) are detected by RM and CMs
- there is no need of correction factors or terms for CMs.

Table 1 provides an overall summary of results of each trials for each instrument included (Thermo R&P FDMS series 8500 version b for PM_{10} and $PM_{2.5}$, Environnement SA MP101M-RST for PM_{10}):

Table 1: Summary of results for French equivalence trials (PM_{10} 24h values)

Candidate Instrument	Site location & type – time period	Equivalence Criteria Met?	Correction Required ?
PM_{10} FDMS series 8500 ver.b (Thermo R&P – 12 min cycle measurement)	Paris suburbs (UB), Marseille city center (UB) } cold period	Meets criteria	No correction required
$PM_{2.5}$ FDMS series 8500 ver.b (Thermo R&P – 12 min cycle measurement)	Paris suburbs (UB), Marseille city center (UB) } cold period	Meets criteria	No correction required
Betâ gauge PM_{10} MP101M-RST (Environnement SA – 24h step time measurement)	Paris suburbs (UB), Marseille city center (UB) } cold period; Aarschot (Belgium - PU), Monterotondo (Italy – PU) } hot period	Meets criteria	No correction required

Conclusions

Results for the two automatic monitoring methods show that the following meet the equivalence criteria set out: TEOM retrofitted with FDMS (for PM_{10} and $PM_{2.5}$); and beta gauge MP101M-RST (for PM_{10}). All these units meet the equivalence criteria set down without the application of correction for slope and/or intercept.
Due to the variability of test sites (in time and space) involving different composition of ambient air and meteorological conditions, it can be assumed that equivalence for equipment tested under the used configuration is valid anywhere else in France under ambient conditions.

Acknowledgement

The authors acknowledge the manufacturers and the Belgian (VMM – ISSeP – IBGE BIM) & Italian (CNR-IIA) institutions for their technical support. This work was supported by the French Ministry of Environment (MEDD) and by the French Agency for Environment and Energy Management (ADEME) in the framework of LCSQA activities.

References

[1] Council Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air

[2] « Demonstration of equivalence of ambient air monitoring methods » - Report by an EC Working group on Guidance for the Demonstration of Equivalence (November 2005)

[3] Technical Report « Evaluation of automatic analyzer MP101M.C for mass concentration measurement of PM10 particulate matter upon request by Environnement SA using reference procedures required by Ministerial Decree n°60 – 2 april 2002 – in application of Directives 1999/30/CE – 2000/69/CE » - 17/11/2004

[4] Standard EN 14907 (2005) « Ambient air quality - Standard gravimetric measurement method for the determination of the $PM_{2.5}$ mass fraction of suspended particulate matter »

EQUIVALENCE OF DIFFUSIVE SAMPLERS TO REFERENCE METHODS: THE OZONE RADIAL DIFFUSIVE SAMPLER AND MEMBRANE-CLOSED PALMES TUBE EXAMPLES

Michel Gerboles[a*], Daniela Buzica[a], Hervé Plaisance[b], A. Piechocki[b], G. Galloo[b]
[a]DG-JRC, Institute for Environment and Sustainability, Joint Research Centre Via E. Fermi, I-21020 Ispra VA, Italy
[b]Ecole des Mines de Douai, Département Chimie et Environnement, F - 59508 Douai

*Tel: +39 0332 785652, fax: +39 0332 789931. michel.gerboles@jrc.it.

Abstract

The study of the equivalence of the NO_2 membrane-closed Palmes tube (MCPT) and of the O_3 Radiello sampler to the reference methods is presented. For NO_2, the equivalence to the annual limit value as stated in the 1st daughter Directive (40 µg m^{-3}) and the reference period of one year was considered. In the case of O_3, the aim was to demonstrate the equivalence for the reference period of 8 hours, at the target value stated in the 3rd daughter Directive which is 120 µg m^{-3}. This study showed that for NO_2, it would be possible to demonstrate the equivalence of the MCPT, but further tests are needed. On the other hand, the radial sampler for O_3 fails to meet the DQO for continuous monitoring.

Introduction

The diffusive sampling [1,2] is a widely used indicative method, which can provide an average concentration of a gaseous ambient air pollutant over a chosen exposure period. Diffusive samplers are devices capable of taking samples of gases from the atmosphere at a rate controlled by a physical process such as gaseous diffusion through a static air layer or a porous material and/or permeation through a membrane, but which does not involve active movement of air through the device[3]. The diffusive sampling method is being increasingly used because it provides a simple and inexpensive alternative to automatic monitoring methods for the determination of air pollution levels and in particular for screening the air quality in agglomeration or larger areas [4,5,6].

The chemiluminescence method [7] is the reference method of the European Directive[8] for monitoring nitrogen dioxide (NO_2). Palmes tubes[9] are used as an indicative method of measurement according to the European Directive. Monitoring of O_3 is generally carried out using continuous measurement techniques like the ultraviolet-photometry[10] which is the reference method of the European Directive[11]. However, these methods are expensive and they require continuous site maintenance, calibration and electric supply. On the contrary, there is a great interest in the development of a cheap and unobtrusive alternative technique such as the diffusive samplers for the measurements of 8-hr O_3 concentrations without the need for electric supply or site calibration
However, the 1st and 3rd European Directives state that a Member State may use any other method apart from the reference method provided that the equivalence of their candidate method to the reference method can be demonstrated. An Equivalent Method, as defined in the new guide for demonstrating equivalence[12], is a method meeting the Data Quality Objectives (DQO) for continuous or fixed measurements as specified in the relevant air quality directives.
Hereafter, the equivalence of the NO_2 Palmes tube and of O_3 Radiello sampler will be evaluated. For NO_2, the equivalence to the annual limit value over one year given in the 1st daughter Directive[**Error! Bookmark not defined.**] (40 µg m^{-3}) will be considered. The DQO is 15 %, meaning that the relative expanded uncertainty should be lower than 15 %. The reference method for NO_2 monitoring is the chemiluminescence method. The candidate method is the membrane-closed Palmes diffusion tube (MCPT) [13] exposed for 2 weeks with a time series covering 1 year (90 %) while the scope of equivalence is based on a semi-rural site in Ispra.

In the case of O_3, the aim is to demonstrate the equivalence to the target value over 8 hours stated in the 3rd daughter Directive for ozone, which is 120 µg m^{-3}. The DQO is 15% and the reference method is the UV photometry. The candidate method will be the O_3 radial diffusive sampler (Radiello®)[14] and the scope of equivalence is based on 5 urban- suburban and rural sites in Italy and France.

Description of samplers

O_3 Radiello sampler

The radial diffusive sampler (see
Figure 1,a) consists of a micro porous polyethylene cylinder of 16 mm in diameter, 50 mm long (47 mm available to the diffusion), 1.75 mm wall thickness and 20-30 µm pore size. Two cellulose acetate caps are soldered with an epoxy adhesive to the cylinder ends. An absorbing cartridge is inserted into the cylinder. This cartridge is filled with silica gel coated with 1,2-bis (4-pyridyl)ethylene (DPE). For exposition, the Radiello sampler is screwed on a plane cellulose acetate equilateral triangle equipped with an attaching clip. All ready-to-use radial diffusive sampler components are commercially available to the Fondazione Salvatore Maugeri (www.Radiello.com). Ambient O_3 diffuses through the porous membrane until the cartridge where it is trapped by reaction with DPE. The diffusion of O_3 molecules is controlled by the coefficient of molecular diffusion of O_3 in air, the geometry of the sampler, the effective area of the pores in the membrane and the gradient between O_3 concentration in ambient air and at the cartridge area where DPE keeps the O_3 concentration close to zero. The concentration of pollutant in air is calculated using the following equation:

$$C_{O_3\,\mu g/m^3} = \frac{v.\left(m_{\mu g} - m_{b\,\mu g}\right)}{d.U_{m^3/min}.t_{min}} = \frac{48/107.11 \; 1.177/0.82\left(m - m_b\right)}{0.996\,U.t} \qquad \textbf{Eq. 1}$$

where v is the coefficient of conversion of PA into O_3: the ratio of O_3 and PA molecular masses (48/107.11 g) divided by the stoechiometric coefficient of O_3 in DPE (0.82) and multiplied by 1.177 (the ratio of the slopes of the PA calibration lines with DPE/without DPE in the calibration standards); m and m_b are the PA masses measured in the exposed cartridge and its blank; d is the desorption efficiency (99.6 %), U is the O_3 uptake rate and t is the exposure time. The uptake rate for O_3 is calculated using equation 2[17]:

$$U_{cm^3.min^{-1}} = 12.2\ln\left(m - m_b\right) + 9.7 + 1.68T_{^\circ C} + 1.14RH_{\%} - 0.042T_{^\circ C}RH_{\%} \qquad \textbf{Eq. 2}$$

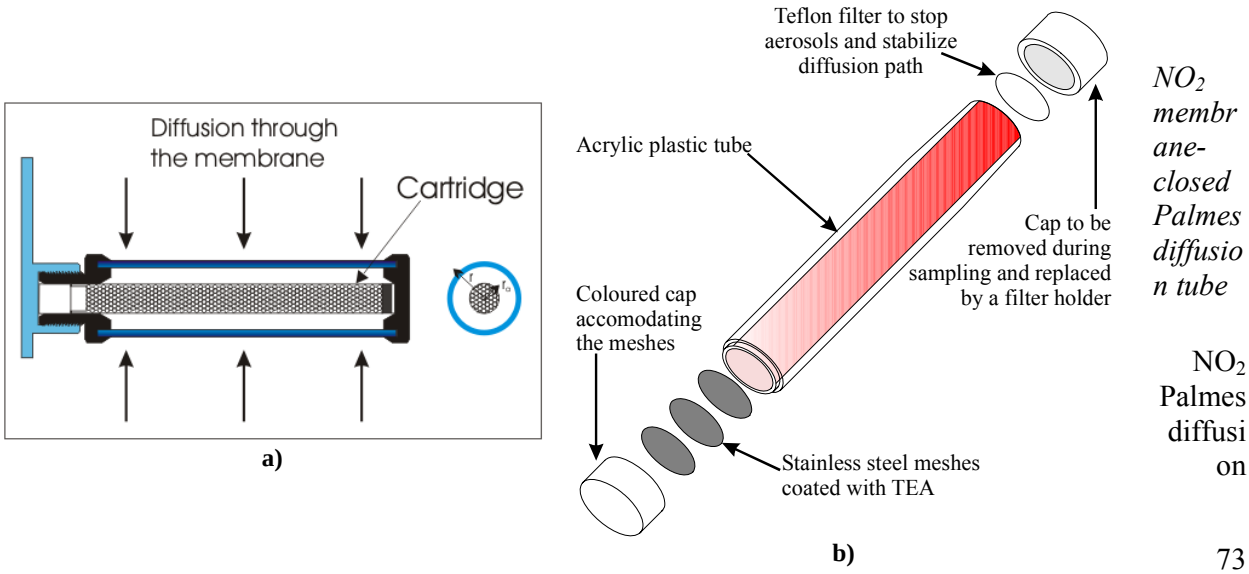


Figure 1 Description of the sampler a) Radiello diffusive sampler; b) Palmes tube

tubes consists of an acrylic tube open at one end and stainless steel meshes coated with TEA at the closed end (see
Figure 1, b). NO_2 Palmes diffusion tubes are commercially available at Gradko International (UK)[15]. NO_2 diffuses through the air in the tube and is trapped as nitrite ion on TEA. The suggested method of preparation of membrane-closed Palmes diffusion tubes (MCPTs) is to clean tubes (Gradko DIF100), membranes (XDIF500F) and caps (XDIFCAP-001, XDIFCAP-003 and XDIFCAP-011) in an orbital shaker using Millipore water and changing the water every half an hour for 3 hours. All samplers are then placed in an oven, at 45 ^{0}C until they are completely dry. The stainless steel mesh discs (XDISC) are cleaned in an ultrasonic bath, at 60 ^{0}C for 5 hours, changing the water every half an hour. Then, they are placed in an oven, flushed with nitrogen, at 125 ^{0}C until they are completely dry. Three clean and dry discs are placed in the coloured cap and 40 µl of a 10% v/v solution of TEA in deionised water is spread all over the meshes using a micropipette. A tube is placed immediately on the coloured cap while the other end is sealed immediately with a membrane for immediate use. It is advised to check if the membrane is correctly placed to make sure that NO_2 diffuses only through the membrane. MCPTs are stored in a plastic container (Gradko Int. Ltd. XDIF30585) in a refrigerator. The NO_2 concentration is calculated using by using an equation derived from Fick's first law:

$$NO_{2\,ppb} = \frac{l}{D\,S}\frac{m}{t} = \frac{m_{ng}}{U_{ng.ppb^{-1}\min^{-1}}\;t_{\min}} \qquad \textbf{Eq. 3}$$

where C is the NO_2 concentration at the temperature and pressure conditions of exposure, m is the mass uptake of nitrite, t is the averaging time, and U is the so-called uptake rate.
The uptake rate [13] is calculated using the following equation, where RH is the relative humidity in %, T is the temperature in ºC, w in the wind speed in m s^{-1}:

$$U = \frac{(134 + 0.86\,RH - 0.00130\,t + 1.28\,T + 3.18\,w)10^{-5}}{2} + \frac{\sqrt{(134 + 0.86RH - 0.00130\,t + 1.28\,T + 3.18\,w)^2\,10^{-10} - 0.64.10^{-5}\,\frac{m}{t}}}{2} \qquad \textbf{Eq. 4}$$

Methodology of demonstration of equivalence

The procedure of demonstration of equivalence for diffusive samplers is given in the Guide of 'Demonstration of equivalence of ambient air monitoring methods'. One has to follow Programme 1 (Manual methods for gases and vapours), laboratory and field test programme for diffusive sampling (Test Programme 1B). First, the uncertainty is estimated based on laboratory experiments. Second, the uncertainty of measurements by the samplers is estimated for the field experiments. The candidate method is said to be equivalent to the reference method if the field uncertainty is smaller than the laboratory one or if the field uncertainty is between the laboratory uncertainty and the DQO. In that case, the Candidate Method is accepted conditionally and further evaluation of the uncertainty from the laboratory tests should be performed.
For the laboratory tests, the reduced test programme with two extreme exposure conditions was considered not sufficient in order to better understand the parameters that affect the sampler. Therefore, an extended laboratory test programme was set up for both NO_2 and O_3 to evaluate the uncertainties of the uptake rate models. The laboratory experiments were performed in an exposure chamber able to control the exposure environment while ensuring homogeneous conditions for all the samplers [16].

The next step is the performance of the field tests. According to the Guide, they shall be performed in such way that the Candidate Method and a Standard Method based on the Reference Method are compared side-by-side. Test sites shall be representative of typical conditions for which equivalence will be claimed. A minimum of 40 measurement results for the Candidate Method per comparison shall be collected.
Regarding uncertainty estimation, first, the between-sampler uncertainty w_{bs} is calculated from the difference of results of the candidate sampler operated in parallel using equation 5 where y_{ij} are the results of measurement j for a single period i, $\overline{y_i}$ is the mean result for period i and p is the number of replicates for period i (for replicate samplers with p > 2). The w_{bs} between sampler uncertainty shall comply with the criteria given in the Guide, i.e. for NO_2 and for O_3 shall be less than 5%.

$$w_{bs}^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\left(y_{i,j} - \overline{y_i}\right)^2}{n(p-1)\overline{y}^2} \qquad \textbf{Eq. 5}$$

Second, one has to calculate the expanded uncertainty of the candidate method. It is assumed that the relationship between measurement results of both method (candidate and reference) can be described by a linear relation as shows equation 6 where x is the average result of the Reference Method and y the one of the candidate method. The expanded uncertainty is calculated using equation 7 where RSS is the sum of residuals resulting from an orthogonal regression given in equation 6, the $u(x_i)$ is the random uncertainty of the reference method. RSS is calculated using equation 8 and the last term is the bias of the candidate method at the limit value/target value x_i .The algorithm of estimation values for b_0 and b_1, and their uncertainty is given in the Guide for the demonstration of equivalence.

$$\hat{y} = b_0 + b_1 x \qquad \textbf{Eq. 6}$$

$$U(y)_{\%} = k.u_c(y) = k\sqrt{\frac{\left(RSS/n-2\right) - u^2(x_i) + \left[b_0 + (b_1 - 1)x_i\right]^2}{y^2}} \qquad \textbf{Eq. 7}$$

$$RSS = \sum\left(y_i - b_0 - b_1 x_i\right)^2 \qquad \textbf{Eq. 8}$$

Demonstration of equivalence

O_3 laboratory experiments

During the laboratory experiments, O_3 was monitored in the exposure chamber using a UV photometry analyser (Thermo Environment model 49C). This analyser was daily calibrated with an O_3 generator SYCOS KT-O3M of ANZYCO (G) certified against a long path UV photometer (UMEG - G).
Details about the determination of O_3 can be found in Plaisance et al.[17]. First, the uptake rate of the Radiello sampler was estimated at several O_3 levels between 20 and 390 µg m^{-3} using sets of 10 samplers exposed for 8 h at 23 ± 1 °C , 43± 2 % of relative humidity and 1 ± 0.2 m s^{-1} of wind speed. The results of these experiments were used to set up an empirical relation between the uptake rate U and the mass of PA in µg, m_{PA}. Subsequently, O_3, averaging time and wind velocity were kept constant. Four experiments were carried out at 23/35 °C and 20/70 % of relative humidity with 6 replicates. By applying two-way analysis of variance (ANOVA, Einax, 1995) on the calculated uptake

rate, temperature, humidity and their interaction were shown to have a significant influence on the O_3 uptake rate. When temperature was high, the uptake rate decreased with the level of relative humidity while for low temperature, the uptake rate increased with the level of relative humidity. Equation 2 gives the uptake rate according to temperature, humidity and temperature-humidity interaction.

Once the uptake rate has been calculated, the next step is the calculation of the uncertainty of the laboratory results. Applying the method given in the GUM[20], the combined uncertainty can be estimated using eq. 9 that is established by derivation of equation 1.

$$u_C^2 = \left(\frac{m-m_b}{d\,U\,t}\right)^2 \left(u_v^2\right) + \left(\frac{v}{d\,U\,t}\right)^2 \left(u_m^2 + u_{m_b}^2\right) + \left(\frac{v(m-m_b)}{d^2 U\,t}\right)^2 u_d^2 + \left(\frac{v(m-m_b)}{dU^2 t}\right)^2 u_U^2 + \left(\frac{v(m-m_b)}{dUt^2}\right)^2 u_t^2 \quad \textbf{Eq. 9}$$

The standard uncertainty for m was estimated using the standard deviation of all sets of 6 replicate samplers. The standard uncertainty for the blanks was 0.17 µg. The major contribution to the combined uncertainty was given by the uptake rate U. In order to calculate this, two experiments in the exposure chamber for 8 hours performed during an inter-comparison exercise[18] were taken into account. The bigger difference between the concentration estimated by the model and the reference value was 19%. We considered a rectangular distribution using this difference and the standard deviation of the uptake rate. Therefore, the uncertainty for the uptake rate was found to be 11%. The standard uncertainty for t was estimated assuming an error of 5 minutes. The expanded uncertainty of the laboratory results were found to be 20% which is not in line with the requirement of the Guide for demonstration of equivalence since it should be less than 15%.

O_3 field programme

A field comparison between the Radiello sampler and UV analyser was carried out at five stations. Four stations belonged to the air quality monitoring network in France and the last one was an EMEP station in Italy. One station was located in the north of France Valenciennes (VA), two in the south: Montpellier (MO) and Martigues (MA) and one in the east: Strasbourg (ST) while the EMEP station was in Ispra (northern Italy) in a pre-alpine region. These sites were selected because they are located in areas favourable to the occurrence of O_3 episodes (two "urban background" sites: MA and VA, two "suburban" sites: MO and ST and one rural site between the Po valley and the Alps). They were also distinct for their meteorological conditions. The UV-photometric analysers in France were regularly calibrated (maximum quarterly) with working standards (portable O_3 generators) which were themselves certified against the national reference standard (a NIST UV long path photometer) while the one in Italy was monthly calibrated using an O_3 portable generator certified against an UMEG long path UV photometer. In France, the measurements were performed during the hot wave period of summer 2003 from 07/28/2003 to 09/16/2003. These uncommon conditions allowed the evaluation of the diffusive sampler for exposure of 8 hours in a wide range of O_3 from 43 to 196 µg.m^{-3} and at high temperatures between 17 to 35°C. The measurements at the EMEP station were carried out in summer 1999 with temperature between 19.1 and 26.7 ºC and O_3 between 58 and 158 µg.m^{-3}. For each sampling, one blank plus one sampler were exposed from 10 h to 18 h. 65 pairs of measurements were obtained.

The relative between–sampler uncertainties w_{bs} was calculated according to eq. 5 and was found to be 4.4% at 120 µg.m^{-3}. This figure satisfies the requirement of 5% given in the Guide of Demonstration of Equivalence in the case of O_3.

The orthogonal regression based method was used to evaluate the field uncertainty. The model-predicted uptake rate given by equation 4 was applied to calculate the concentrations measured by diffusive sampler The orthogonal regression line (equation 6) for the UV-photometry (x) versus the

radial samplers (y) values had a slope of $b_1 = 1.08 \pm 0.04$ and an intercept $b_0 = -9.0 \pm 5.0$. Equation 7 was used to calculate the uncertainty of field experiments. The square root of the sum of residuals gave 12.2 µg m^{-3}. Following the recommendation of the Guide, the random uncertainty of the standard method $u(x_i)$ was set to 5 % of the candidate method at a target value (120 $\mu g.m^{-3}$) and the bias a the target value was 0.3 $\mu g.m^{-3}$. Consequently, the estimation of the relative expanded uncertainties (**Table 1**) of an individual 8-hour measurement U(y) gave 20 % (see **Table 1**). The trend of relative expanded uncertainty versus O_3 concentration is given in **Figure 2**.

Table 1: Relative expanded uncertainty U(y) of the Radiello sampler calculated with the modelled uptake rate at the target limit of the ozone Directive (120 $\mu g.m^{-3}$). The table gives the slope and intercept with their standard uncertainty of the regression line, then the combined uncertainty $u_c(y)$ calculated as the sum of the residuals of the orthogonal regression line, of the random uncertainty of the UV photometric method and the bias between the values of the Radiello sampler and the UV photometric method.

	Modelled uptake rate
Slope $b_1 \pm u(b_1)$	1.08 ± 0.04
Intercept $b_0 \pm u(b_0)$	-9.0 ± 5.0
$\sqrt{RSS/n-2}$, Sum of residuals of the orthogonal regression	12.2
$u(x_i)$, random uncertainty of the reference method (5 %)	3
Bias at LV $\mu g.m^{-3}$, $b_0 + (b_1 - 1).x$	-0.3
$u_c(y)$, combined uncertainty $\mu g.m^{-3}$	10.2
U(y) Relative expanded uncertainty %	20

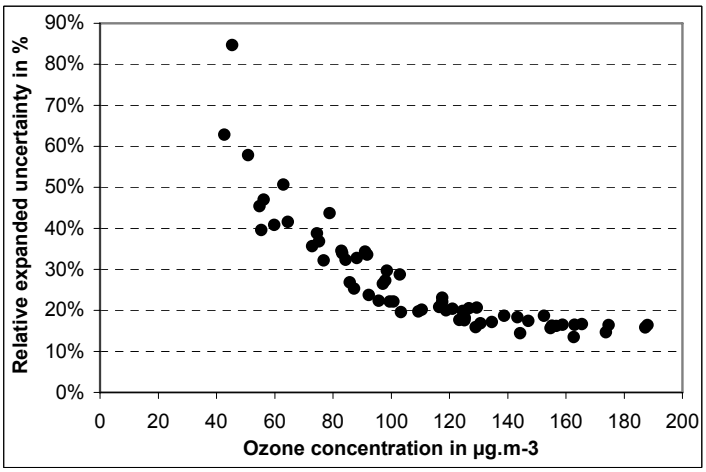


Figure 2: Relative expanded uncertainty versus ozone using the modelled uptake rate.

NO_2 laboratory tests

The chosen approach for the determination of uptake rate and its uncertainty was to conduct a fractional factorial plan of experiment for the study of 5 factors, in which the factors, each at two levels, were combined according to Hadamard's matrix. This design required 16 experiments and enabled a statistical analysis of the significant factors affecting the uptake rate for the different combination of conditions. The NO_2 reference value in the exposure chamber was fully traceable to SI

units by use of a primary reference gas mixture for generation and monitoring of NO_2 concentrations in the exposure chamber. The NO_2 concentrations measured by the continuous analyser were corrected for the interference of humidity and for the drift, assuming linear drift between calibration adjustments[19]. The corrections were applied for all measurements performed.
The results of the laboratory tests led to a model equation able to estimate the uptake rate**Error! Bookmark not defined.**. This model equation showed to have a 2 %-dependence on NO_2. NO_2 in the model equation is therefore substituted by $m.U^{-1}.t^{-1}$. The model equation then becomes a second order polynomial in U which could be solved using its positive root. The positive root becomes the new model equation that can predict the uptake rate of a specific exposure (it was given in equation 4).

The next step is to calculate the uncertainty of the laboratory results. The basic method is provided by the GUM [20] Applying this method, the combined uncertainty can be estimated using equation 9 that is established by derivation of Eq. 1.

$$u_C^2 = \left(\frac{1}{U\,t}\right)^2 \left(u_m^2\right) + \left(\frac{m}{U^2\,t}\right)^2 u_U^2 + \left(\frac{m}{U\,t^2}\right)^2 u_t^2 + 2\left(\left(\frac{m}{U^3\,t^2}\right) u_m\; u_U\; r(m,U)\right) \quad \textbf{Eq. 10}$$

The standard uncertainty for m (0.027 µg) was estimated using the pooled mean standard deviation of all sets of 6 replicate samplers of the factorial plan of experiment (0.023 µg) and of blanks (0.015 µg). The major contribution to the combined uncertainty was given by the uptake rate U. It was estimated by the standard deviation of the residuals (0.00013 ng ppb^{-1} min^{-1}) between the observed and estimated uptake rate for all trials of the factorial plan of experiment. The standard uncertainty for t was estimated assuming an error of 30 minutes. The correlation between the uptake rate and mass of nitrite absorbed by the sampler was equal to –0.78 while the one between the uptake rate and averaging time was null. The expanded uncertainty of the laboratory results was found to be 10% and it respects the DQO criteria of less than 15% as specified in the Guide of Demonstration of equivalence.

NO_2 field programme

For field tests programme, the monitoring station (AIRMON) in Ispra (I) was chosen. The location of the site was semi-rural. The AIRMON station was accredited under ISO 17025 for the measurement of several pollutants including nitrogen oxides. Meteorological information was also available. Sets of 6 samplers were exposed every 2 weeks. The range of concentration was varying between 8 and 45 µg m^3. According to the Guide[12**Error! Bookmark not defined.**], at least 20% of the results should have been greater than or equal to the upper assessment threshold specified in the Daughter Directive. This condition was not met in this case. 40 pairs of data (passive samplers - chemiluminescence) were collected (see Gerboles et al.[13**Error! Bookmark not defined.**] for the full dataset).
The relative between – sampler uncertainties w_{bs} is calculated from the differences of results of the candidate sampler operated in parallel according to eq. 5. The requirement for between – sampler/instrument uncertainty for NO_2 (5%) is satisfied in the case of Palmes sampler (3.7 %).
Instead of the orthogonal regression, the method of ISO 6143 [21] was used to estimate the field uncertainty of NO_2 diffusive samplers measurements. The uncertainty of the chemiluminescence values were needed to apply this method. Gerboles et al.[19] gave full details of the estimation of the uncertainty of chemiluminescence measurements and this method was used hereafter. The regression line (equation 6) obtained by application of the method of ISO 6143 for the chemiluminescence (x) versus the MCPT (y) values had a slope of $b_1 = 1.05 \pm 0.045$ and an intercept $b_0 = -0.24 \pm 0.93$ with covariance $cov(b_0,b_1) = -0.037$. Then the combined uncertainty $u_c(x_c)$ of a MCPT value was calculated using eq. 10 where the three first terms were derived from the GUM (the sampler uncertainty determined with the laboratory experiments, the uncertainty of slope and intercept, the covariance) and the last one shows the contribution of the bias between the chemiluminescence and the MCPT values.

The uncertainty is calculated for only one exposed sampler since w_{bs} is added. The relative expanded uncertainty for field measurements with MCPT, $U(x_c)$, is given in **Figure 3.**

$$u_c^2(x_c) = b_1^2 u_c^2(y_c) + u^2(b_0) + y_c^2 u^2(b_1) + 2y_c \operatorname{cov}(b_0, b_1) + w_{bs}^2 + \left(b_0 + \left(b_1 - 1\right)y_c\right)^2 \qquad \textbf{Eq. 11}$$

where y_c is the NO_2 concentration determined using eq. 3, $u_c(y_c)$ is its uncertainty determined using Eq.10.

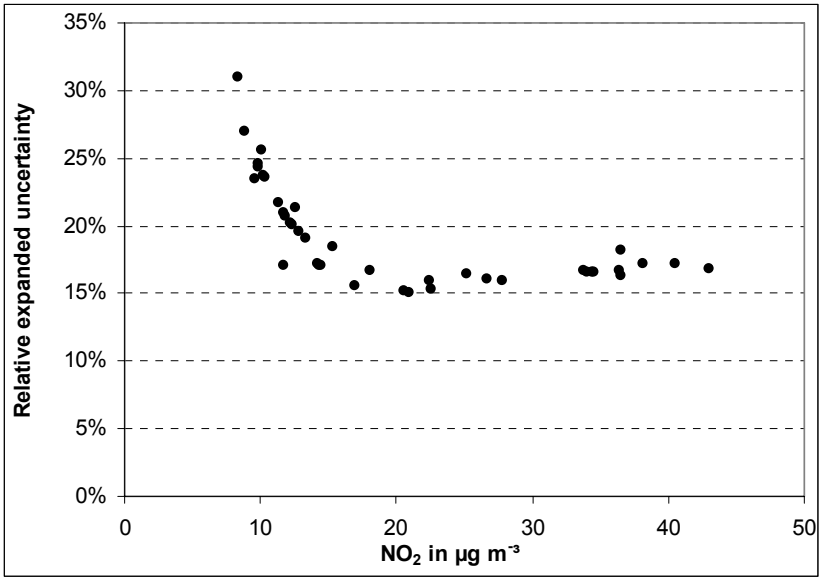


Figure 3: Relative expanded uncertainty for NO_2 measured with diffusive samplers during the field tests

Discussion and Conclusion

According to the laboratory tests, the uncertainty of the MCPT (10%) is smaller than the DQO. During the field tests, the relative expanded uncertainty decreased towards less than 15% for NO_2 higher than 20 µg m^{-3} which is the DQO for continuous measurements of the Directive. For concentrations higher than 10 µg m^{-3}, the relative expanded uncertainty was still lower than 25%, the DQO for indicative measurement. More studies are necessary to complete the data set with NO_2 values higher than the limit value.
However, the equivalence should be demonstrated for exposure period of 1 year with 90 % of coverage corresponding to 23 periods of 2-week samplings. Therefore the uncertainty of the yearly average is decreased compared to the 2-week measurements, in particular random contributions will be lowered by the square root of twenty three [22] . In equation 11, w_{bs} and $u(y_c)$ are affected by this decrease since both of them becomes random errors when a time series is considered. Our estimation of the resulting relative expanded uncertainty gave a value of 12% at the limit value of 40 µg.m^{-3}. Therefore the DQO is reached, a fact that is confirmed by the results of the field experiments in Ispra. In 1999, 2000 and 2001 the averages by the chemiluminescence method was 18.9 µg.m^{-3}, 27.3 µg.m^{-3} and 18.8 µg.m^{-3} while the MCPTs gave 17.8 µg.m^{-3} (n = 11), 25.7 µg.m^{-3} (n = 11) and 20.4 µg.m^{-3} (n = 18).

For O_3 however, the DQO for continuous monitoring is 15% while the uncertainty in the laboratory was found to be 20%. The uncertainty in the field was also 20%. Since the uncertainty in the field is bigger then the DQO, the radial diffusive sampler fails to meet the DQO for continuous monitoring.

Nevertheless, the Radiello sampler for O_3 is meeting the DQO for indicative measurements defined in the European Directive as its expanded uncertainty is far lower than 30 % around the target value of 120µg.m^{-3} and for 8-h average concentration.

References

[1] Brown, R.H., The use of diffusive samplers for monitoring of ambient air, IUPAC, 1993.
[2] Hangartner, M., Kirchner, M., Werner, H., Evaluation of passive methods for measuring ozone in the European Alps, The Analyst, 1996, 121, 1269-1272
[3] European Committee for Standardization, Ambient air quality - Diffusive samplers for the determination of concentrations of gases and vapours - Requirements and test methods - Part 1: General Requirement EN 13528-2, Brussels, 2002.
[4] Gerboles, M., Amantini, L., Detimmermann, F., Paryssat, M., Studio della distribuzione del biossido di azoto a Bologna mediante campionatori passivi, Office for Official Publications of the European Communities, Luxembourg, 1998, EUR 18145 IT.
[5] Stevenson, K., Bush, T., Mooney, D., Five years of nitrogen dioxide measurement with diffusion tube samplers at over 1000 sites in the UK, Atmospheric Environment, 2001, 35, 281-287.
[6] Ferm, M., Rodhe, H., Measurements of Air Concentrations of SO2, NO2 and NH3 at rural and remote sites in Asia, Journal of Atmospheric Chemistry, 1997, 27, 17-29.
[7] European Committee for Standardization, Ambient air quality-Measurement method for the determination of the concentration of nitrogen dioxide and nitrogen monoxide by chemiluminescence, EN 14211, Brussels, 2005.
[8] Council Directive 1999/30/CE of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air, Official Journal of European Communities, L 163/41.
[9] Palmes, E.D., Gunnison, A. F., Personal monitoring device for gaseous contaminants, 1973, American Industrial Hygiene Association Journal, 78-81.
[10] European Standard, 2005, EN 14625, Ambient air quality - Standard method for the measurement of the concentration of ozone by ultraviolet photometry, Brussels, Belgium
[11] Directive 2002/3/EC of the European Parliament and the Council of 12 February 2002 relating to ozone in ambient air, Official Journal of the European Communities, L 67/14
[12] European Commission, 2005, Demonstration of Equivalence of ambient air monitoring methods, Report by an EC Working group, http://europa.eu.int/comm/environment/air/pdf/equivalence_report3.pdf
[13] M. Gerboles, D. Buzica and L. Amantini, Modification of the Palmes diffusion tube and semi-empirical modelling of the uptake rate for monitoring nitrogen dioxide, Atmos. Environ., 2005, 39, 2595–2608.
[14] Fondazione Salvatore Maugeri, Instruction manual for Radiello sampler, version 1/2003, http://www.radiello.com.
[15] http://www.gradko.co.uk/introduction.shtml
[16] Buzica, D., Gerboles, M., Amantini, L., Laboratory and field inter-comparisons of NO2 diffusive samplers, Office for Official Publications of the European Communities, Luxembourg, 2003, EUR 20860 EN.
[17] H. Plaisance, M. Gerboles, A. Piechocki, F. Detimmerman, E. de Saeger, Radial diffusive sampler for the determination of 8-h ambient ozone concentrations, Environmental Pollution 148 (2007) 1- 9
[18] Gerboles, M., Buzica, D., Amantini, L., Lagler, F, Laboratory and field comparison of measurements using the available diffusive samplers for ozone and nitrogen dioxide in ambient air, Journal of Environmental Monitoring, 2006, 8, 112 - 119
[19] Gerboles, M,. Lagler, F., Rembges, D. , Brun, C., 2003, Assessment of uncertainty of NO2 measurements by the chemiluminescence method and discussion of the quality objective of the NO2 European Directive, Journal of Environmental Monitoring, 5, 529–540.

[20] International Organisation for Standardisation, 1995, Guide to the expression of uncertainty in measurement, ISBN 92-67-10188-9, ISO, Geneva
[21] International Organisation for Standardisation (ISO) 2001. Gas analysis—determination of the composition of calibration gas mixtures—comparison methods. ISO 6143:2001 (E), ISO, Geneva.
[22] International Organisation for Standardisation (ISO), 2002, Air quality – Determination of the uncertainty of the time average of air quality measurements ISO 11222:2002 (E), Geneva.

PM10 EQUIVALENCE TESTS IN SWITZERLAND: RELATIONSHIP BETWEEN A TEOM-FDMS AND THE REFERENCE METHOD AT A SUBURBAN SITE

Ch. Hueglin, R. Gehrig, Beat Schwarzenbach

Empa, Swiss Federal Laboratories for Materials Testing and Research, CH-8600 Duebendorf, christoph.hueglin@empa.ch, Tel. +41 44 823 46 54

Introduction

The present work summarizes the results of equivalence tests of an automated PM10 monitor and the reference method according to a guideline for demonstration of equivalence of ambient air monitoring methods [1] as performed in Switzerland by Empa. Automated PM10 monitors (TEOM-FDMS 8500, Thermo Electron Corporation, East Greenbush NY, USA) were tested in the field for equivalence with a manual gravimetric method fulfilling the equivalence criteria of EN12341 (PM10 collection using a high volume sampler, Digitel DA-80H). The tests were performed at three different sites, a suburban site (Duebendorf), a rural site located at the kerb of a motorway (Haerkingen), and a rural site at elevated altitude (Rigi, 1040m asl).

Method

The "Field test programme" for particulate matter described in [1] has been used to test the equivalence of the TEOM FDMS compared to the European reference method.

At the suburban site in Duebendorf, two TEOM-FDMS instruments were run collocated from August 26 2005 to December 11 2006. The data were recorded on a hourly basis and aggregated to daily values. During the same measurement period, a high volume sampler (HVS) equipped with glass fibre filters (HVS/GF, Ederol 227/1/60) was operated at a total of n=102 days. In addition, a second HVS equipped with quartz fibre filter (HVS/QF, Whatman QMA) was run on n=109 days during the January 1 2006 to December 11 2006. At the two other sites Haerkingen and Rigi, daily PM10 values from TEOM-FDMS instruments and HVS are available for every fourth day from February 10 2007 to September 1 2007 (n=48) and December 19 2006 to September 1 2007 (n=63), respectively.

The collocated TEOM FDMS measurements in Duebendorf allowed to calculate the between sampler uncertainty of these instruments. Between sampler uncertainties of the HVS equipped with the two used filter types was determined based on collocated 24-hourPM10 measurements from September 27 2005 to November 26 2005 (n=31, glass fibre filter Ederol 227/1/60) and a second campaign from March 3 2006 to April 4 2006 (n=31, quartz fibre filter Whatman QMA).

Results and discussion

Figure 1 shows the time series of collocated daily PM10 measurements using two TEOM FDMS instruments at the suburban site in Duebendorf (August 26 2005 to December 11 2006). The daily PM10 values obtained with the gravimetric method using high volume samplers equipped with glass fibre filters (Ederol 227/1/60) and quartz fibre filters (Whatman QMA) are also included in **Figure 1**.

During the first few months of parallel TEOM FDMS measurements, the differences of the instrument readings were relatively large. The noise level of one of the instruments was suspiciously high, the

microbalance was therefore replaced on January 9 2006. Since then, the agreement between the two instruments was much improved. However, all available data were used for the equivalence test procedure. The calculated between sampler uncertainty for the candidate method (CM - here the TEOM FDMS) is u_{bs}=1.99µg/m^3 (n=474). When all data before January 9 2006 are excluded, a clearly reduced between sampler uncertainty u_{bs}=1.28µg/m^3 (n=337) is obtained. The corresponding values for the reference method (RM) are lower, for glass fibre (GF) and quartz fibre filters (QF) between sampler uncertainties of 0.92 µg/m^3 and 0.59 µg/m^3 were found, respectively.

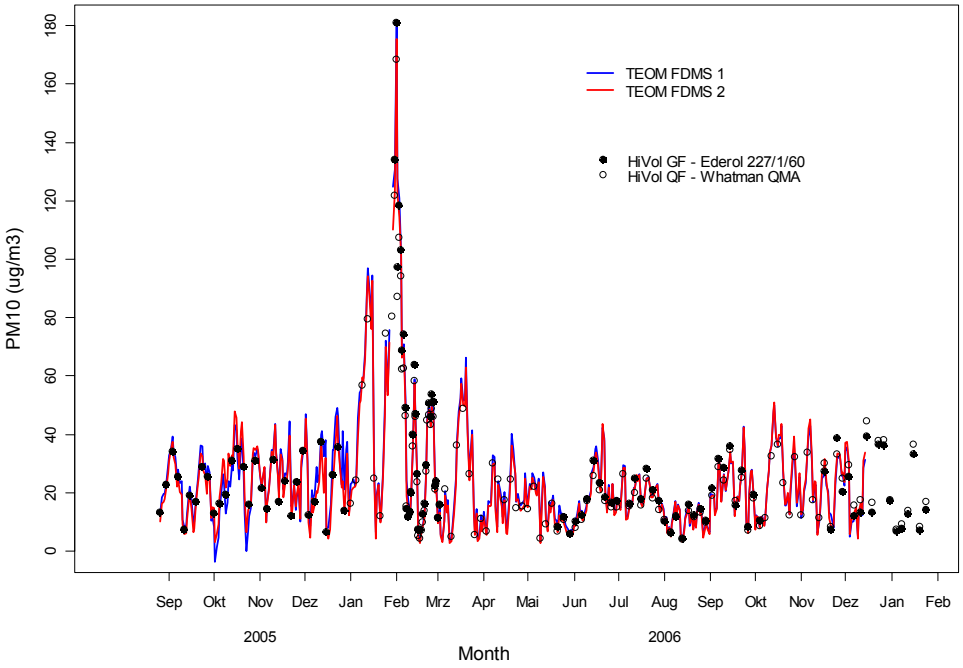


Figure 1: Daily PM10 measurements with two TEOM FDMS instruments and two high volume samplers equipped with two different filter types at the suburban site in Duebendorf, Switzerland, from August 26 2005 to December 11 2006.

The comparison of the two TEOM FDMS and the HVS equipped with glass fibre filters was excellent (**Figure 2**), the TEOM FDMS passed the equivalence test (performed with the provided Excel-Template) without correction (**Table 1**). The daily PM10 values used for this analysis ranged from 4.2 to 181.0 µg/m^3. Exceptionally high PM10 levels occurred in January and the beginning of February 2006, when stable high-pressure weather was leading to temperature inversions and accumulation of air pollutants within the boundary layer in large parts of Switzerland and the neighbouring countries. During this high air pollution episode, the HVS equipped with quartz fibre filters showed substantial losses of PM10. This can be concluded from the fact that during this time period the PM10 values measured with the HVS/QF were systematically lower than with the HVS/GF and the two TEOM FDMS instruments (**Figure 1**). Consequently, orthogonal regression analyses of the two TEOM FDMS versus the HVS equipped with quartz fibre filters yielded slopes and intercepts that are significantly larger than one and significantly different from zero, respectively (**Figure 3**). Slope and intercept corrections were therefore necessary to pass the equivalence test. The observed losses of PM10 from quartz fibre filters and the noticed implication on the equivalence test are unsatisfactory.

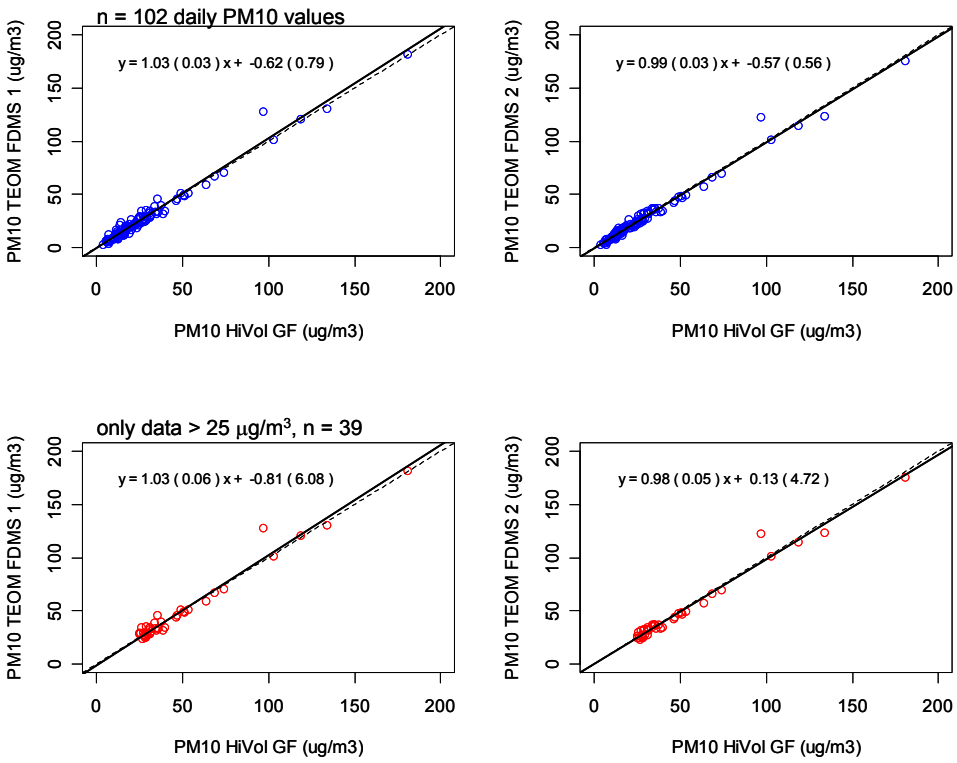


Figure 2: Scatter-plots for daily PM10 of two collocated TEOM FDMS instruments versus a manual gravimetric method fulfilling the equivalence criteria of EN12341 (high volume sampler equipped with glass fibre filter Ederol 227/1/60).

When measurements during the high air pollution episode are rejected, the found linear relationship between TEOM FDMS and HVS/QF is CM=1.03(± 0.01)·RM - 0.22(± 0.31), which is similar to the TEOM FDMS and HVS/GF comparison (**Table 1**). However, the slope is significantly different from 1, therefore a slope correction is necessary here to meet the equivalence criteria. It is noteworthy to mention, that the small bias of the TEOM FDMS as expressed by the slope that is slightly larger than 1 appears here because the random noise of the instrument was very small. A larger random noise would increase the confidence interval for the slope estimation, consequently a similar instrument but with a larger random noise would have passed the equivalence test without correction.

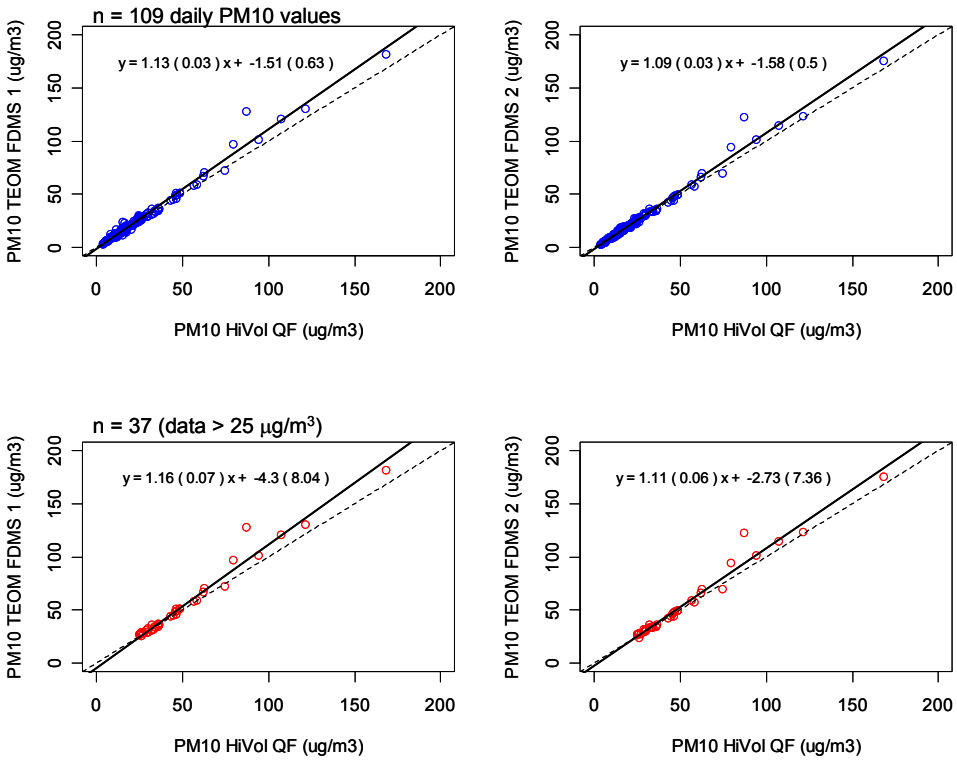


Figure 3: Scatter-plots for daily PM10 of two collocated TEOM FDMS instruments versus a manual gravimetric method fulfilling the equivalence criteria of EN12341 (high volume sampler equipped with quartz fibre filter Whatman QMA).

Field tests of TEOM FDMS instruments have also been performed at two rural sites (Haerkingen and Rigi) and have recently been started at an urban background site. The Haerkingen site is located on the Swiss plateau at an altitude of 430m above sea level and 20 m north of a four-lane motorway with heavy traffic (55'000-80'000 vehicles per day). Here, the TEOM FDMS passed the equivalence test without correction factor (**Table 1**). In contrast, a slope correction factor was necessary to achieve equivalence at the Rigi site, which is located at the foothills of the Alps at 1040m above sea level. The obtained relationship between TEOM FDMS and the HVS equipped with quartz fibre filters is CM=1.14(± 0.05)·RM - 0.06(± 0.21), the reason for this systematic difference is currently unknown. In order to gain a better understanding of differences between the TEOM FDMS and the gravimetric method using HVS, the measurements will be continued at all of these sites and subsets of the data from individual sites will be analysed separately (e.g. different seasons, temperature, humidity).

Table 1: Summary of the performed tests for equivalence of TEOM FDMS instruments with the EN reference method for particulate matter.

Site	**Duebendorf**	**Duebendorf**	**Duebendorf**	**Haerkingen**	**Rigi**
Type	Suburban	Suburban	Suburban	Rural, traffic	Rural, above 1000m asl
Period	Aug 2005 – Dec 2006	Jan 2006 – Dec 2006	Feb 2006 – Dec 2006[a]	Feb 2007 – Aug 2007	Dec 2006 – Aug 2007
PM Metric	PM10	PM10	PM10	PM10	PM10
Filter type used for RM	GF Ederol 227/1/60	QF Whatman QMA	QF Whatman QMA	QF Whatman QMA	QF Whatman QMA
CM	TEOM FDMS	TEOM FDMS	TEOM FDMS	TEOM FDMS	TEOM FDMS
Nr of data pairs	102	109	94	48	61
Mean of RM data	27.9µg/m^3	26.8µg/m^3	19.3µg/m^3	20.2µg/m^3	11.1µg/m^3
Range of RM data	4.2 – 181.0µg/m^3	3.9 – 168.3µg/m^3	3.9 – 58.3µg/m^3	7.2 – 48.5µg/m^3	0.7 – 53.4µg/m^3
Between-RM uncertainty	0.92µg/m^3	0.59µg/m^3	0.59µg/m^3	0.59µg/m^3	0.59µg/m^3
Between-CM uncertainty	1.99µg/m^3	1.99µg/m^3	1.99µg/m^3	1.99µg/m^3	1.99µg/m^3
Relation between RM and CM	CM=1.01RM-0.56	CM=1.11RM-1.54	CM=1.03RM-0.22	CM=1.00RM +0.30	CM=1.14RM +0.05
Relative uncertainty at Limit value	7.5%	6.92%	3.11	4.50	3.67
Equivalence? with/without correction?	Yes, without corr.	Yes, with S+I corr.	Yes, with S corr.	Yes, without corr.	Yes, with S corr.

[a] when few outliers (reference method) are rejected, see text.

Conclusions

PM10 field measurements at three sites in Switzerland during all seasons indicate that TEOM FDMS monitors can be accepted as being equivalent to the reference method. It seems that no correction factors are needed at some sites, while correction of the TEOM FDMS data is necessary at others. More data are needed to improve the understanding about possible differences between the TEOM FDMS and the reference method.

References

[1] DEMONSTRATION OF EQUIVALENCE OF AMBIENT AIR MONITORING METHODS, Report by an EC Working group on Guidance for the Demonstration of equivalence, November 2005.

NON SEASONAL VARIABILITY AND THE EFFECT OF LOGGING AVERAGES ON TEOM AND FDMS MEASURMENTS MADE DURING THE UNITED KINGDOM PM_{10} and $PM_{2.5}$ EQUIVALENCE TRIALS.

David Harrison, Jeff Booker and Richard Maggs. Bureau Veritas HS&E Ltd., 30 Great Guildford Street, London SE1 0ES, United Kingdom.
Email: david.harrison@uk.bureauveritas.com; Tel: +44 (0)207 9026170.

Introduction

The United Kingdom PM Equivalence Trials [1] took place from late 2004 to early 2006 across four sites for each of two seasons. These trials included comparisons of a number of PM_{10} candidate samplers (TEOM, FDMS, Met-One BAM and Opsis SM200) against the EN12341 filter reference method (PM_{10} KFG). The $PM_{2.5}$ FDMS candidate method was compared to the EN14907 filter reference method ($PM_{2.5}$ Leckel).

This paper highlights two areas of research that have come about since the publication of the report in June 2006:

- Previous studies in Switzerland [2] and the United States [3,4] have shown that the ratio of the TEOM and filter based measurement varies seasonally with ambient temperature. The UK Equivalence Study showed that for the PM_{10} TEOM there were no slope or intercept correction factors that could be employed to make the expanded uncertainty at the limit value (WCM) less than 25 % for all data sets. The results of the UK equivalence study were reanalysed using ambient temperature as a variable to see if any improvement in the expanded uncertainty could be derived.

- TEOM data in the UK network are logged using data loggers connected to the analogue outputs of the instrument. A temporal resolution of 15 minute averaging is used. In contrast, FDMS units (which are found to be equivalent to both EN12341 and EN14907) report PM data through a 6 minute switching cycle. The implications for connecting an FDMS to the same logging system as the TEOM are discussed.

The Effect of ambient temperature on TEOM concentrations

The TEOM data from the UK Equivalence Trials were reanalysed using ambient temperature as a variable. The TEOM data studied are referred to as TEOM(0,1,1), and are the 24 hour average TEOM concentration without any internal correction factors (typically Slope = 1.03 and Intercept = 3), or external correction factor (Typically Slope = 1.3).

The ratio of PM_{10} KFG / TEOM (0,1,1) was plotted against ambient temperature (**Figure 1**). Current EU Working Group Guidance on Equivalence [5] allows for the data to be corrected by the slope and intercept of the 'All Data' dataset. In this scenario the correction factor applied is: $Y = X(1.59-0.015T)$. The results of applying this correction factor can be observed in **Figure 2** and **Table 1**.

As the expanded uncertainty (W_{CM} / %) at both the Annual and Daily Limit Values is, in the majority of instances greater than 25 %, a single correction method for TEOMs incorporating ambient temperature would not be suitable for deployment in the United Kingdom.

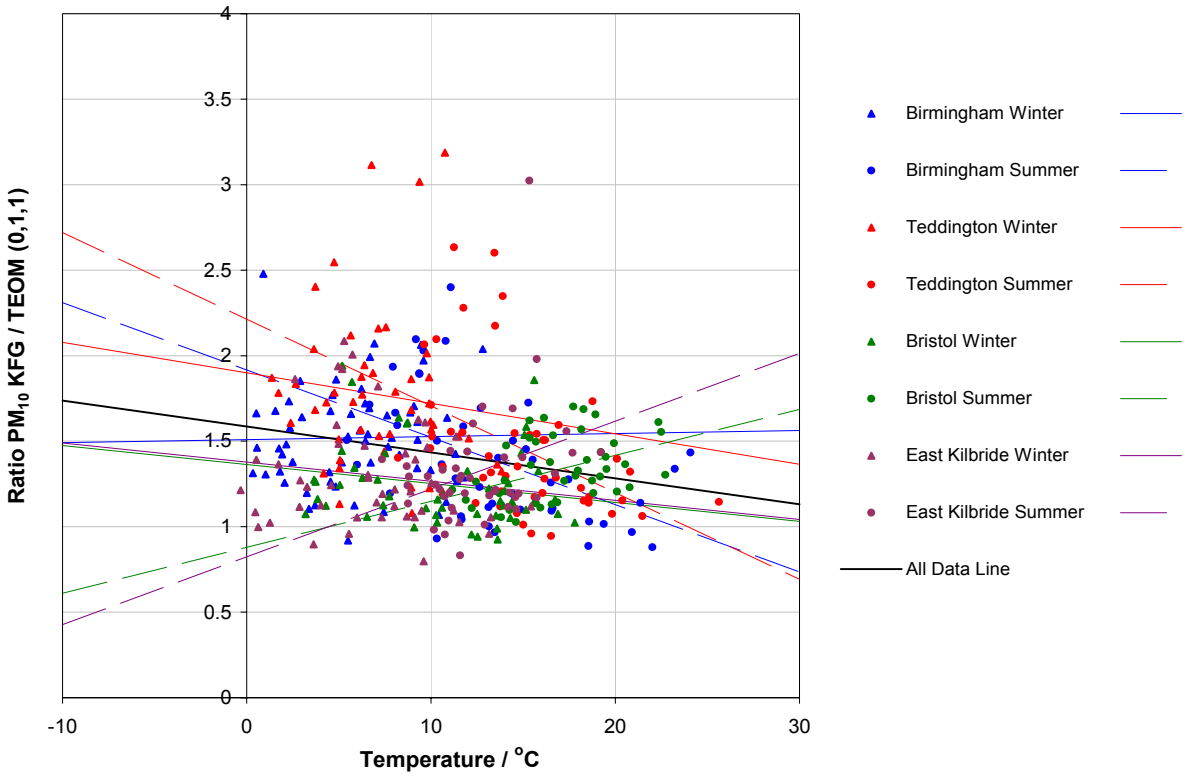


Figure 1: The Variation of the ratio of PM_{10} KFG/ TEOM (0,1,1) versus Ambient Temperature for each of four sites over two seasons.

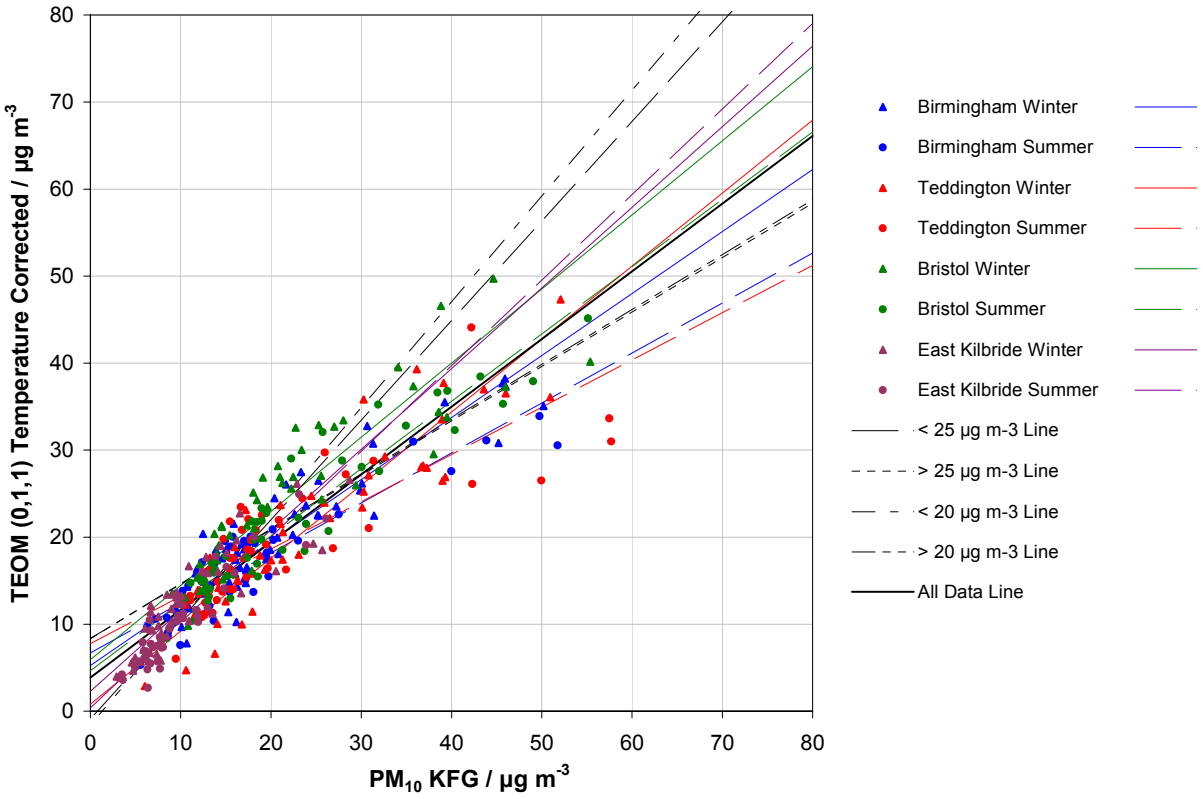


Figure 2: The Variation of the TEOM(0,1,1) data corrected by Y = X(1.59-0.015T) against the PM_{10} KFG for each of four sites over two seasons.

Table 1: The Variation of the TEOM(0,1,1) data corrected by Y = X(1.59-0.015T) against the PM_{10} KFG for each of four sites over two seasons. Green indicates that the criteria are within the specified limits, whereas red indicates that it is outside.

TEOM (0,1,1) Temperature Corrected	Dataset	1 hour		24 hour		Orthogonal Regression				Annual Limit Value of 40 µg m^{-3}		Daily Limit Value of 50 µg m^{-3}		
		n_{bs}	u_{bs}	n_{bs}	u_{bs}	n_{c-s}	r^2	Slope (b) +/- u_b	Intercept (a) +/- u_a	W_{CM} / %	% > 50 % LV	W_{CM} / %	% > 50 % LV	(n_{ES},n_{EC})
Individual Campaigns	Birmingham Winter	2600	0.55	107	0.33	59	0.850	0.712 +/- 0.038	5.261 +/- 0.865	33.81	39.0	37.94	25.4	(1,0)
	Birmingham Summer	2080	0.85	85	0.50	45	0.842	0.574 +/- 0.036	6.690 +/- 0.745	52.96	20.0	59.10	13.3	(1,0)
	Teddington Winter	2938	1.08	122	0.48	47	0.846	0.839 +/- 0.050	0.822 +/- 1.358	34.23	53.2	32.89	40.4	(2,0)
	Teddington Summer	2218	1.67	76	1.19	48	0.645	0.543 +/- 0.053	7.771 +/- 1.248	56.33	27.1	62.43	20.8	(2,0)
	Bristol Summer	1351	1.82	53	0.46	44	0.886	0.774 +/- 0.041	4.655 +/- 1.054	26.29	43.2	29.03	31.8	(1,0)
	Bristol Winter	2182	1.34	89	1.00	49	0.806	0.852 +/- 0.056	5.921 +/- 1.353	21.18	44.9	17.95	30.6	(1,0)
	East Kilbride Summer	1554	0.60	61	0.41	44	0.871	0.983 +/- 0.054	0.375 +/- 0.609	8.12	6.8	6.65	2.3	(0,0)
	East Kilbride Winter	1721	0.81	66	0.96	47	0.761	0.927 +/- 0.068	2.294 +/- 0.835	14.14	10.6	12.30	4.3	(0,0)
All Campaigns	All Data	16644	1.15	659	0.72	383	0.813	0.778 +/- 0.018	3.866 +/- 0.386	31.34	31.1	32.60	21.4	(8,0)
Annual Limit Value of 40 µg m^{-3}	< 20 µg m^{-3}	13535	0.89	264	0.53	264	0.696	1.215 +/- 0.040	-1.574 +/- 0.537	37.72	-	-	-	-
	> 20 µg m^{-3}	3109	1.91	121	0.60	119	0.527	0.630 +/- 0.045	8.355 +/- 1.523	40.68	-	-	-	-
Daily Limit Value of 50 µg m^{-3}	< 25 µg m^{-3}	14946	0.95	301	0.54	301	0.733	1.145 +/- 0.034	-0.913 +/- 0.495	-	-	27.79	-	-
	> 25 µg m^{-3}	1698	2.25	84	0.60	82	0.387	0.625 +/- 0.065	8.332 +/- 2.452	-	-	46.85	-	-

- Denotes Not Applicable.

Theoretical discussion on the effect of logging averages on TEOM and FDMS

In the UK, multi-instrument pollution monitoring stations conventionally use a site logger to collect data from all instruments. TEOMs are currently being upgraded through retro-fitting of FDMS units in light of the outcome of the PM Equivalence Programme. There follows a discussion on the implications for connecting the FDMS to the same logging system as a TEOM. In the UK Equivalence Programme, hourly data were collected directly from the FDMS, and as such, are not affected by the following discussion. Similarly, all FDMSs now installed in the UK are also logged directly from the instrument data buffer, and as such no problems are anticipated.

TEOMs are typically logged as 15 minute averages using the site logger connected to the analogue outputs of the instrument. The four preceding 15 minute averages are automatically averaged to create an hourly average concentration which is disseminated, and on which 24-hour average data are subsequently derived.

The FDMS switches between measuring the mass change of the oscillating filter in filtered and particle laden air every 6 minutes. The voltage output is switched every 6 minutes to give an average of the 1 hour period prior to the current 6 minute period.

If the FDMS were to be connected to the site logger, and the data polled and collated in an identical manner to the TEOM, this could potentially introduce a bias to the hourly and 24 hour averages (**Figure 3**). The effect can best be observed for the one hour data. The red area corresponds to the data that one wants to record (i.e. from midnight to 1 AM), whereas the blue area corresponds to the area that would be logged if the FDMS were connected to the data logger. The pyramid shape results from logging a rolling average, and leads to data from a 2 hour period being logged (from 11 PM to 1 AM), rather than the desired 1 hour.

In addition to this effect, our experience from the UK equivalence study suggests the gain of mass in the sampling phase is greater than the loss of mass in the filtered phase. As such, there is the potential for bias to the measurements depending on which of the two phases occurs first in the hourly average.

Further, if the logger and FDMS clocks are not synchronized, this will further complicate the data capture.

If the ambient PM_{10} concentration were variable, this could lead to a significant skew in the data, and it is possible that it would be significant for 24 hour averages at the limit value, thereby leading to errors in the reported number of exceedences when compared to gravimetric reference methods. No hourly Limit Values for PM_{10} or $PM_{2.5}$ exist. As such, the potential for errors induced by averaging are less significant; however in the interests of data validity, every effort should be made to limit such errors.

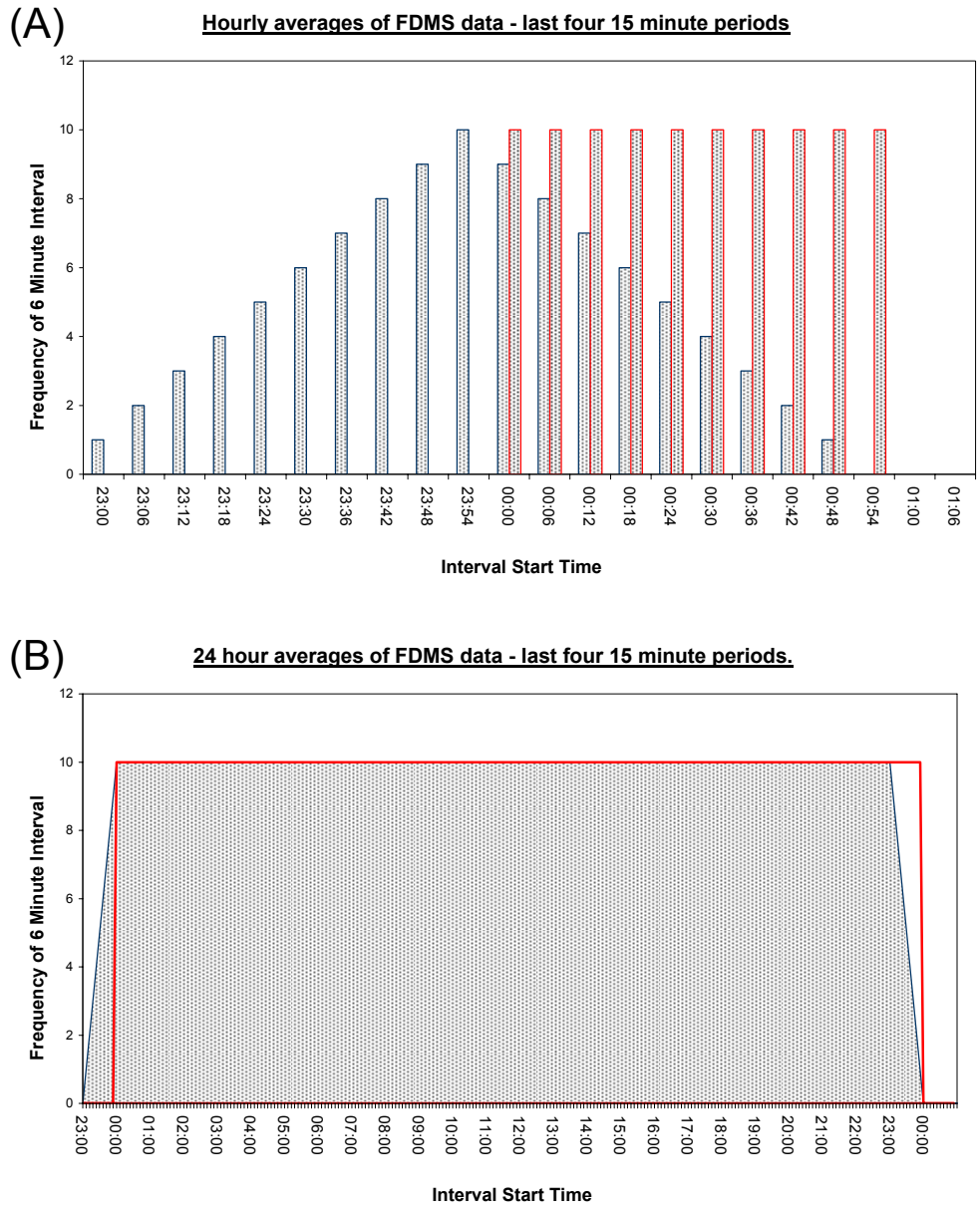


Figure 3: Frequency of 6 minute interval in (A) 1 hour and (B) 24 hour average of FDMS data when connected to a data logger averaging 15 minute periods. Blue corresponds to actual averages, and Red to the average required for the data to be accurate.

This discussion indicates that the sum of 15 minute analogue output averages cannot be assumed to give the same results where external logging facilities are employed.

Future studies

It is thought that the ambient temperature effects observed in the US studies are following the seasonal variability of volatile species such as ammonium nitrate in the North-eastern United States [6]. In the United Kingdom, ammonium nitrate concentrations are known to vary regionally, and be highest in the South East corner of Britain [7]. Concentrations may also vary significantly according to whether the Air Mass originates from continental Europe, or the Atlantic Ocean. Results show that a single correction factor incorporating ambient temperature would not be suitable for use in the United Kingdom. However, it may be possible to utilise different temperature based correction factors in different areas of the United Kingdom to reflect the occurrence of differential levels of ammonium nitrate, and other volatile species.

Further work to quantify the potential bias introduced through external logging systems for the FDMS is required. Studies took place in the UK from June to September 2007. Instruments were logged using a digital logging system operating at 1 minute sampling frequency. These data will be compared against the hourly data obtained direct from the instrument. The results will indicate whether the logger provides the same hourly concentration as that measured by the FDMS, and also whether the 12 minute FDMS averages can be used to recalculate the hourly average correctly. These data will then in turn be used to validate the theoretical study presented in this paper.

Conclusions

- A single correction method for TEOMs incorporating ambient temperature would not be suitable for deployment in the United Kingdom as the expanded uncertainty at both the Annual and Daily Limit Values is consistently greater than 25 %.
- Attaching an FDMS to the same data logging system as a TEOM may introduce bias to reported PM concentrations obtained from the instrument for hourly measurements. The impact of this bias on reported 24-hour data is required to be quantified.

References

[1] UK Equivalence Programme for Monitoring of Particulate Matter; David Harrison (2006); http://www.airquality.co.uk/archive/reports/cat05/0606130952_UKPMEquivalence.pdf

[2] PM10-Umrechnungsmodelle für TEOM und betameter-messreihen; Jürg Heldstab and Marx Stampfli (2005). INFRA and Stampfli Mathematics Publication.

[3] Continuous $PM_{2.5}$ Method and Data Issues: Insuring regional consistency in real-time PM data; George Allen (2004). http://www.marama.org/calendar/events/presentations/2004_10Mon/Allen-regiona_pm_consistency-issues.pdf

[4] New York State Urban and Rural Measurements of Continuous $PM_{2.5}$ Mass by FDMS TEOM and BAM: Evaluation and Comparisons with the FRM; Dirk Felton, Oliver Rattigan, Kenneth Demerjian and James Schwab (2005) http://www.marama.org/calendar/events/presentations/CM05-March/DFelton_Presentation.pdf

[5] Demonstration of Equivalence of Ambient Air Monitoring Methods (Version 3).
URL: http://europa.eu.int/comm/environment/air/pdf/equivalence_report3.pdf

[6] George Allen, NESCAUM, personal communication.

[7] www.airquality.co.uk

ANNEX

GUIDELINE TO DEMONSTRATION OF EQUIVALENCE OF AMBIENT AIR MONITORING METHODS

DEMONSTRATION OF EQUIVALENCE OF AMBIENT AIR MONITORING METHODS

Report by an EC Working group on
Guidance for the Demonstration of Equivalence

1 INTRODUCTION 4

2 REFERENCES TO NORMS 5

3 TERMS, DEFINITIONS AND ABBREVIATIONS 6

3.1 TERMS AND DEFINITIONS 6
3.2 ABBREVIATIONS 6

4 DEFINITION OF EQUIVALENCE 7

5 PROCEDURE FOR DEMONSTRATION OF EQUIVALENCE 8

5.1 FLOW SCHEME 8
5.2 GENERAL 10
5.3 REQUIREMENTS FOR LABORATORIES 10
5.4 SCOPE OF EQUIVALENCE 11
5.4.1 Limiting conditions 11
5.4.3 Extent of tests required 12
5.5 PRACTICAL APPROACH TO EQUIVALENCE TESTING 14
5.6 OPERATION OF THE CANDIDATE METHOD 15

6 SELECTING A TEST PROGRAMME 15

6.1 GENERAL 15
6.2 MEASUREMENT METHODOLOGY 16
6.3 MEASUREMENT TRACEABILITY 16
6.4 SPECIFICATION OF TEST PROGRAMMES 16

7 TEST PROGRAMME 1 - MANUAL METHODS FOR GASES AND VAPOURS 18

7.1 GENERAL 18
7.2 OVERVIEW OF THE TEST PROCEDURES 18
7.3 LABORATORY TEST PROGRAMME 19
7.3.1 Test programme 1A: pumped sampling 20
7.3.2 Test Programme 1B. Diffusive sampling 27
7.4 FIELD TEST PROGRAMME 29
7.4.1 General 29
7.4.2 Experimental conditions 29
7.4.3 Evaluation of the field test data 30
7.4.4 Evaluation of results of field tests 32

8 TEST PROGRAMME 2: AUTOMATED MEASUREMENT SYSTEMS FOR GASES 33

8.1 GENERAL 33
8.2 OVERVIEW OF THE TEST PROCEDURES 33
8.3 DEFINITIONS APPLICABLE TO AUTOMATIC MEASUREMENT SYSTEMS 35
8.4 LABORATORY TESTS 35
8.4.1 Test concentrations 35
8.4.2 Response time 35
8.4.3 Short–term drift 36
8.4.4 Repeatability for continuous measuring CMs 37
8.4.5 Carry over and repeatability for CMs collecting samples onto a sorbent prior to analysis. 38
8.4.6 Lack of fit (linearity) 38
8.4.7 Differences between sample and calibration port 39
8.4.8 Effect of short-term fluctuations in concentration (averaging test) 40
8.4.9 Variation in sample-gas pressure 41
8.4.10 Variation in sample-gas temperature 42
8.4.11 Surrounding temperature variation 42

8.4.12 Variation due to supply voltage ... *43*
8.4.13 Cross-sensitivity to interfering substances ... *44*
8.4.14 NO_2 converter efficiency ... *47*
8.5 FIELD TEST ... 48
8.5.1 General ... *48*
8.5.2 Experimental conditions ... *48*
8.5.3 Evaluation of data collected ... *48*
8.6 DETERMINATION OF THE COMBINED MEASUREMENT UNCERTAINTY ... 52
8.7 CALCULATION OF THE EXPANDED LABORATORY UNCERTAINTY OF CANDIDATE METHOD ... 52
8.8 EVALUATION OF TEST RESULTS ... 52

9 TEST PROGRAMME 3 – METHODS FOR PARTICULATE MATTER ... 53

9.1 GENERAL ... 53
9.2 OVERVIEW OF THE TEST PROCEDURE ... 53
9.3 LABORATORY TEST PROGRAMME ... 54
9.3.1 General ... *54*
9.3.2 Application of automated filter changers ... *54*
9.3.3 Different weighing conditions ... *54*
9.4 FIELD TEST PROGRAMME ... 55
9.4.1 General ... *55*
9.4.2 Experimental conditions ... *55*
9.5 EVALUATION OF DATA COLLECTED ... 56
9.5.1 Suitability of datasets ... *56*
9.5.2 Calculation of performance characteristics ... *56*
9.5.3 Calculation of the combined uncertainty to be assigned to the Candidate Method ... *58*
9.5.4 Calculation of the expanded uncertainty of Candidate Method ... *58*
9.6 EVALUATION OF RESULTS OF FIELD TESTS ... 58
9.7 APPLICATION OF CORRECTION FACTORS OR TERMS ... 58
9.8 EXAMPLES ... 60

10 TEST PROGRAMME 4 – SPECIATED PARTICULATE MATTER ... 61

10.1 GENERAL ... 61
10.2 OVERVIEW OF THE TEST PROCEDURES ... 61
10.3 LABORATORY TEST PROGRAMME ... 61
10.3.1 General ... *61*
10.3.2 Test programme ... *62*
10.4 FIELD TEST PROGRAMME ... 66
10.4.1 General ... *66*
10.4.2 Experimental conditions ... *67*
10.4.3 Evaluation of test results ... *67*
10.4.4 Evaluation of results of field tests ... *69*

11 REPORTING REQUIREMENTS ... 70

12 REFERENCES ... 71

ANNEX A ... 73

ANNEX B ... 74

ANNEX C ... 75

ANNEX D ... 76

1 INTRODUCTION

One of the objectives of the European Directive 96/62/EC ('Air Quality Framework Directive', ref.1) is to '*assess the ambient air quality in Member States on the basis of common methods and criteria*'.
The European Air Quality Directives ('Daughter' Directives) associated with this Framework Directive, which relate to limit or target values for specified atmospheric pollutants, are Directives 1999/30/EC, 2000/69/EC, 2002/3/EC and 2004/107/EC [2-5]. These specify the principles of the reference methods to be used for the measurement of concentrations of these pollutants. In addition, they specify *Data Quality Objectives* that have to be met for the performance of specific measurement tasks. These data quality objectives include minimum requirements for:

- expanded uncertainties of measurement results at the limit or target value(s) set for each pollutant
- time coverage of the measurements in relation to the reference period of the limit or target values
- data capture when using the measurement method, i.e., effective measurement time.

CEN Technical Committee (TC) 264, the Committee for Air Quality Measurement Methods, has produced, or is producing, Standard Methods to satisfy these directives. CEN TC 264's remit was to ensure these were validated against the data quality objectives given in the relevant Directives. In order to harmonize the approaches of the various ambient air Working Groups, in particular for the assessment of the measurement uncertainties, a CEN Report was prepared in which the principles for these uncertainty assessments are laid down (report CR 14377).
A Member State (MS) when implementing the Directives should use the Standard Methods, but the directives allow Member States the possibility to *'use any other method which it can demonstrate gives results equivalent to the above* (standard) *method'*.

This Report describes the principles and methodologies to be used for the demonstration of the equivalence of alternative methods to the EN Standard Methods. It is intended for use by laboratories nominated by National Competent Authorities (see Directive 96/62/EC [1]) to perform the tests relevant to the demonstration of equivalence of ambient-air measurement methods.
The building blocks of the equivalence demonstration procedure are presented in Figure 1 below.

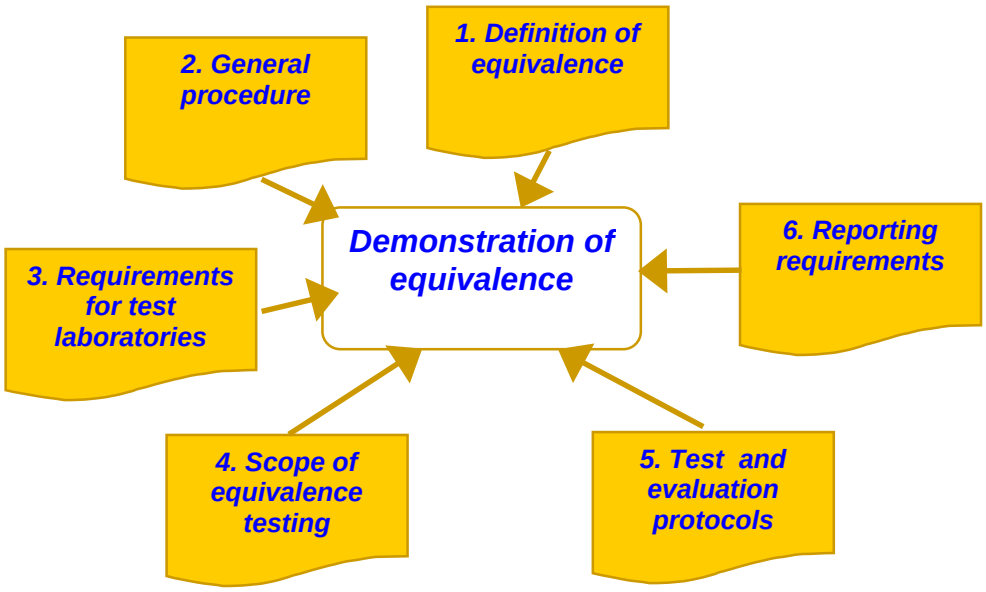


Figure 1. Building blocks for equivalence demonstration

2 REFERENCES TO NORMS

This clause incorporates by dated or undated reference, provisions from other publications. These normative references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to applies.

EN 12341	1998	Air Quality – Determination of the PM_{10} fraction of suspended particulate matter – Reference method and field test procedure to demonstrate reference equivalence of measurements
ENV 13005	1999	Guide to the expression of uncertainty in measurement
EN-ISO 17025	1999	General requirements for the competence of testing and calibration laboratories
CR 14377	2001	Approach to uncertainty estimation for ambient-air measurement methods
EN-ISO 14956	2001	Air quality – Evaluation of the suitability of a measurement method by comparison with a stated measurement uncertainty
EN 13528 pt1	2002	Ambient air quality – Diffusive samplers for the determination of gases and vapours – Requirements and test methods – Part 1: General requirements
EN13528 pt2	2002	Ambient air quality – Diffusive samplers for the determination of gases and vapours – Requirements and test methods – Part 2: Specific requirements and test methods.
EN13528 pt3	2003	Ambient air quality – Diffusive samplers for the determination of gases and vapours – Part 3: Guide to selection, use and maintenance.
ISO 6142	2000	Gas analysis. Preparation of calibration gas mixtures – Gravimetric methods
ISO 6143	2000	Gas analysis. Comparison methods for the determination of calibration gas mixtures
ISO 6144	2002	Gas analysis. Preparation of calibration gas mixtures – Static volumetric methods
ISO 6145		Gas analysis. Preparation of calibration gas mixtures – Dynamic volumetric methods. All Parts

3 TERMS, DEFINITIONS AND ABBREVIATIONS

3.1 *Terms and definitions*

3.1.1	Automated (Measurement) Method/System	A measurement method or system performing measurements or samplings of a specified pollutant in an automated way
3.1.2	Candidate Method	A measurement method proposed as an alternative to the relevant Reference Method for which equivalence has to be demonstrated
3.1.3	Continuous Measurements	Measurements performed at a fixed site on a continuous basis.
3.1.4	Equivalent Method	A method other than the Standard Method for the measurement of a specified air pollutant, that meets the Data Quality Objectives for continuous or fixed measurements specified in the relevant air quality directive.
3.1.5	Fixed Measurements	Measurements performed at a fixed site. [1]
3.1.6	Limit value	A level fixed on the basis of scientific knowledge, with the aim of avoiding, preventing or reducing harmful effects on human health and/or the environment as a whole, to be attained within a given period and not to be exceeded once attained. [1]
3.1.7	Manual (Measurement) Method	A measurement method by which sampling is performed on site, with sample analysis performed in the laboratory.
3.1.8	National Competent Authority	Authority or body designated by a Member State as responsible for the approval of measuring devices (methods, equipment, etc.). [1]
3.1.9	Reference Method	The principle of a method established by EC legislation for the measurement of a specified ambient air pollutant
3.1.10	Standard Method	A method standardized by CEN enacting a Reference Method for a specified ambient air pollutant
3.1.11	Target value	A level fixed with the aim of avoiding more long-term harmful effects on human health and/or the environment as a whole, to be attained where possible over a given period. [1]

3.2 *Abbreviations*

AMS	Automated Measurement System
CM	Candidate Method
CRM	Certified Reference Material
DQO	Data Quality Objective
EC	European Commission
EU	European Union
IR	Infrared
MM	Manual Method
MS	Member State
NCA	National Competent Authority
PM	Particulate Matter
PSM	Primary Standard Material
PT	Proficiency Testing
RM	Reference Method
UV	Ultraviolet

4 DEFINITION OF EQUIVALENCE

There are two different ways of defining an Equivalent Method:

1. A measurement method that meets the requirements set from the viewpoint of fitness-for-purpose for the intended use of the reference method [6]

2. A measurement method giving results that do not differ from those of the Reference Method within a specified statistical uncertainty [7].

The application of one or other definition requires a considerably different approach to the evaluation of equivalence:

The first definition implies that a Candidate Method should fulfil the data quality objectives set for the Reference Method specified in the relevant directive.

The second definition implies that the results obtained with both methods should be determined to be statistically insignificantly different, for example at the 95% confidence level. It does not automatically imply compliance with the uncertainty data quality objectives. In addition, a combination of low 'random' uncertainties of both methods with a systematic difference in results between methods may lead to rejection of a Candidate Method, which according to the first definition would be acceptable.

It should also be noted that, although not specifically stated in the air quality directives or their associated explanatory documents – on the basis of consultations with various experts in the monitoring field – the intended application of the Reference Methods specified in these directives is for continuous or fixed measurements.

Within the framework of air quality measurements, this choice of definition has already been made in the Document specifying 'Terms of Reference for CEN/TC 264 Ambient-air Standards' (see e.g. Report CR 14377 Annex C). These Terms of Reference state that methods other than the Reference Method may be used for the implementation of the directives provided that they fulfil the minimum data quality objectives specified in the relevant directive. Hence, for the current document the principles of first definition apply.

Therefore, considering the intended use of the reference methodologies, the following definition will be used for the demonstration of equivalence:

> *'An Equivalent Method to the Standard Method for the measurement of a specified air pollutant, is a method meeting the Data Quality Objectives for continuous or fixed measurements specified in the relevant air quality directive'*

NOTE 1. The use of the Standard Methods is not restricted to continuous or fixed measurements.
NOTE 2. Where a Candidate Method fails to meet the uncertainty data quality objective of the Standard Method, it may still be able to meet the uncertainty data quality objective for indicative methods.
NOTE 3. For automated measurement systems for gases all relevant uncertainty sources must be assessed and the Candidate Method must pass all the prescribed individual performance criteria, in addition to the overall uncertainty criteria, in order to conform with all the requirements of the relevant EN Norms.
NOTE 4. Equivalence may be granted for regional situations within a Member State, but also for situations encompassing more than one Member State. The latter case offers an incentive for 'inter-Member State' consultation prior to the performance of equivalence testing.

Tables 1a and 1b give an overview of limit or target values, data quality objectives, Reference Methods and EN Standard Methods for compounds under existing or proposed EU legislation.

Table 1a. Limit values, data quality objectives, Reference Methods and EN Standard Methods.

Compound	*Limit value ($\mu g.m^{-3}$)*	*Reference period*	*Data Quality Objective*		*Principles of Reference Method as specified by Daughter Directives*	*EN Standard Method (or working documents)*
			Expanded Uncertainty (%)	*Data Capture (%)*		
Sulphur dioxide	350	1 h	15	90	Ultraviolet-fluorescence	EN 14212
	125	24 h	15	90		
	20	1 y	15	90		
Nitrogen oxides	200 (NO_2)	1 h	15	90	Chemiluminescence	EN 14211
	40 (NO_2)	1 y	15	90		
	30 (NO_x)	1 y	15	90		
Carbon monoxide	10 $mg.m^{-3}$	8 h	15	90	Non-dispersive infrared spectrometry	EN 14666
Benzene	5	1 y	25	45	Pumped sampling + analysis by gas chromatography	EN 14662 parts 1-3
PM_{10}	50	24 h	25	90	PM_{10} reference sampler (EN 12341)	EN 12341
	40	1 y	25	90		
Lead	0,5	1 y	25	90	PM_{10} reference sampler + analysis by atomic spectrometry	EN 14902

Table 1b. Target values, data quality objectives, Reference Methods and EN Standard Methods.

Compound	*Target value ($\mu g.m^{-3}$)*	*Reference period*	*Data Quality Objective*		*Principles of Reference Method as specified by Daughter Directives*	*EN Standard Method (or working documents)*
			Expanded Uncertainty (%)	*Data Capture (%)*		
Ozone	120	8 h	15	90/75	Ultraviolet photometry	EN 14665
Benz[a]-pyrene	1 $ng.m^{-3}$	1 y	50	90	PM_{10} reference sampler + analysis by liquid chromatography – fluorescence or gas chromatography – mass spectrometry	
Arsenic Cadmium Nickel	6 $ng.m^{-3}$ 5 $ng.m^{-3}$ 20 $ng.m^{-3}$	1 y	40	90	PM_{10} reference sampler + analysis by atomic absorption spectrometry or inductively-coupled plasma – mass spectrometry	EN 14902

NOTES

1. Limit/target values are in $\mu g.m^{-3}$ unless otherwise stated, expressed at 20 °C and 101,3 kPa for gases and vapours; for PM, metals and benz[a]pyrene they are expressed at ambient conditions
2. The expanded uncertainty is defined at the 95% confidence level.
3. The uncertainty of the Reference Method, which is derived for a shorter averaging period used during laboratory and field validation trials, applies to the longer averaging times specified in the directives (CR 14377).

5 PROCEDURE FOR DEMONSTRATION OF EQUIVALENCE

5.1 *Flow scheme*

A flow scheme depicting the procedure for equivalence demonstration is given in Figure 2.

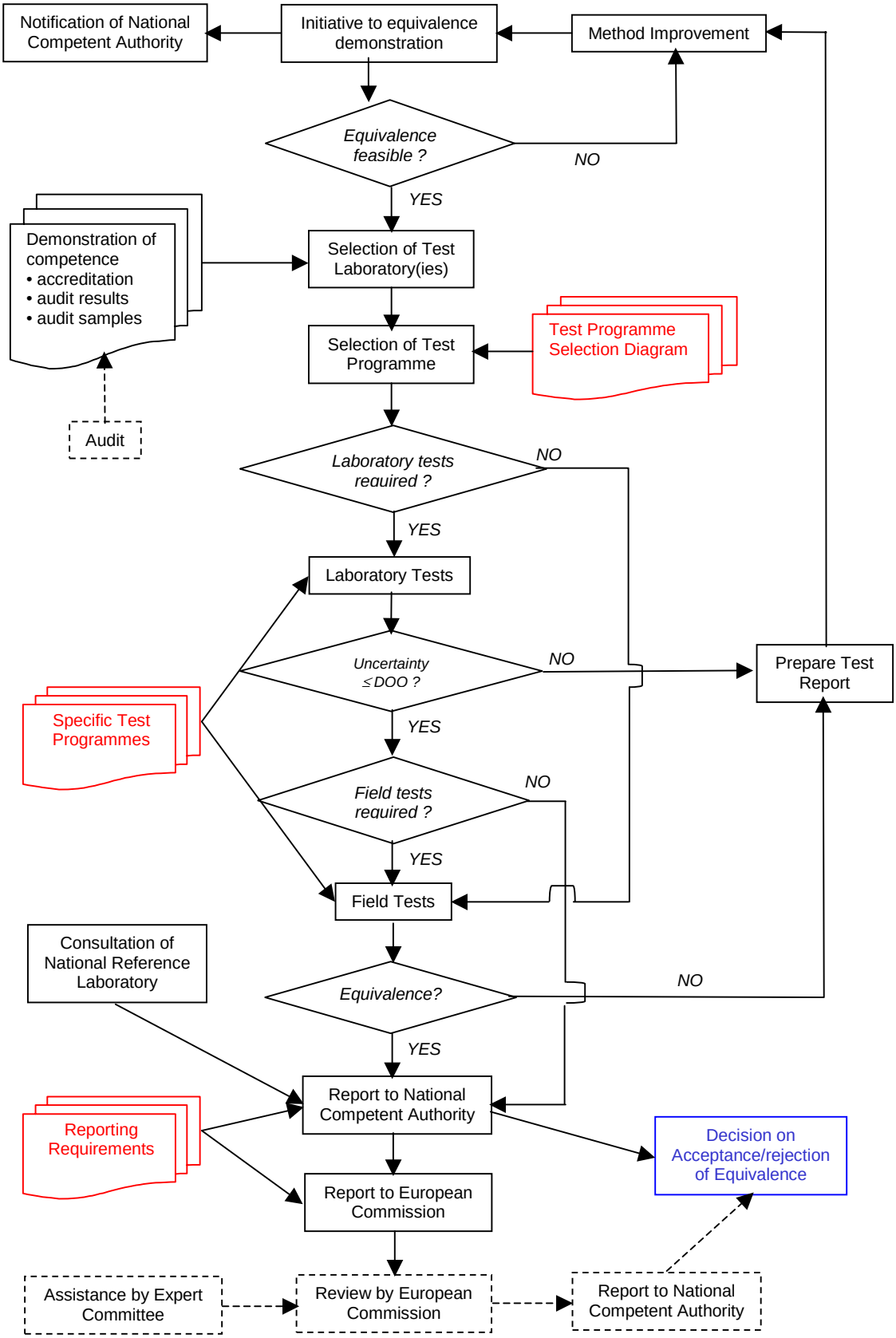


Figure 2. Flow scheme of the procedure for demonstration of equivalence

5.2 General

A Member State may propose methods that deviate from the Reference Method defined in the Daughter Directives [2-5] and elaborated in the EN Standard Methods [8-18] given in Table 1. Consequently, the responsibility for the demonstration of equivalence of the proposed Candidate Method rests with the National Competent Authority. This authority bears responsibility for the quality of national air quality monitoring data. In the process of demonstrating equivalence (see Figure 2) the National Competent Authority (NCA) may delegate its responsibility to a National Reference Laboratory. The NCA remains responsible for the final decision on the acceptance or rejection of a Candidate Method as equivalent to the EN Standard Method, and for reporting to the European Commission.

The initiative for the use of 'equivalent' methods may arise from a National Competent Authority or from a national or regional laboratory performing air quality measurements related to the implementation of the Daughter Directives. In the latter case, the laboratory proposing the use of a method shall notify its National Competent Authority, and perform a preliminary assessment of the Candidate Method in order to ensure that the method:

- fulfils the requirements of data capture and time coverage set for the continuous/fixed measurements, e.g., a Candidate Method for the measurement of concentrations of nitrogen dioxide for comparison with the 1-hour limit value, shall be able to provide a data capture of 90% or more for hourly averaged measurement results, and

- has the potential for meeting the uncertainty requirement for the Standard Method at the limit value concentration.

When the Candidate Method passes this preliminary assessment, the test and evaluation programme relevant to the Candidate Method can be selected using the flow scheme given in Figure 3.

If at any stage of the test programme the measurement uncertainty of the Candidate Method fails to meet the relevant Directive's uncertainty criterion, then the equivalence evaluation may be terminated, and a report of the results obtained prepared for the Competent Authority. This may be used as a basis to reduce relevant uncertainty sources - after which tests appropriate to these uncertainty sources may be repeated, and the resulting uncertainty again compared with the uncertainty criterion.

Following completion of the relevant test and evaluation programme, the results of these tests and evaluations shall be reported to the Competent Authority. The Competent Authority will then decide on the acceptance or rejection of the Candidate Method as an Equivalent Method. In the case of acceptance, an evaluation report with conclusions should be submitted to the European Commission for review. The European Commission in its review may wish to consult a committee of experts about the claim for equivalence.

The European Commission reserves the right to question and to reject the use of a particular method if the equivalence is not sufficiently demonstrated. The methodology presented in this report is an acceptable way of sufficiently demonstrating equivalence. If the Commission produces a negative conclusion on the claim for equivalence, then the Competent Authority should reconsider its decision.

5.3 *Requirements for laboratories*

The laboratories performing the required tests shall be independent of manufacturers or suppliers of equipment used for implementing the Candidate Method.

The laboratories performing the tests necessary for the demonstration of equivalence shall be able to demonstrate technical competence for these tests. These may be the laboratory/laboratories already using the Candidate and/or Standard Method, but may also be

different laboratories, subject to fulfilment of the requirements for laboratories. It is strongly recommended that laboratories work in full compliance with the requirements of EN-ISO 17025, as demonstrated through a formal accreditation.

In the absence of a formal accreditation, compliance with the requirements of EN-ISO 17025 should be demonstrated through an independent audit performed by an auditor with specific experience in the use of the relevant Standard and Candidate Methods. A demonstration of competence by achieving acceptable performance in a suitable Proficiency Testing (PT) scheme is considered useful additional information. In the absence of such a scheme, measurements of a series of appropriate test samples with satisfactory results are strongly recommended for demonstrating competence. Test samples shall be such that the concentration(s) of the compound(s) to be measured is (are) traceable to primary standard materials (PSM) or certified reference materials (CRM).

Note: for the purpose of the supply of suitable test samples, the National Competent Authority may consult an appropriate National Reference Laboratory and/or accreditation body.

5.4 *Scope of equivalence*

5.4.1 Limiting conditions

It is possible for equivalence to be granted for specific 'regional' conditions (the composition of ambient air, meteorological conditions etc). However, in order to promote an economy of scale it is recommended that regional or national laboratories consult others prior to equivalence testing, and cooperate in order to broaden the scope of equivalence. In all cases, however, where the scope of equivalence is restricted in any way, the Equivalent Method should only be applicable over the pollutant concentration range and conditions that were tested for compliance with the relevant Daughter Directive.

In claims to equivalence, limiting conditions shall be specified where relevant. Such limiting conditions should include:

- Composition of the ambient air, i.e., concentration ranges of the specific pollutant and relevant cross interfering species;
- Meteorological conditions, i.e., ranges of temperature, atmospheric humidity, and wind velocity;
- Geographical conditions, such as at specific locations.

5.4.2 Generalization of equivalence claims

For many methods, equivalence that has been proven using the approach described in this Report can be assumed to be valid anywhere else under ambient conditions. Moreover, the test programmes described here generally attempt to demonstrate equivalence for as wide a range of conditions as possible, including practical 'extremes'.
However, this generalization may not hold for all pollutants. This is particularly the case for PM: The semi-volatile fraction, which depends on location and ambient conditions, is not retained in the sample to the same extent by different measuring methods. In addition, current PM levels being close to the limit values, many Member States are required to perform PM measurements throughout their entire territory or in large parts of it, and thus a variety of types of location and ambient conditions are usually involved. Consequently, it may be that equivalence for PM measurements that is established under the conditions described in 9.4.2 of this Report (taking into account where relevant the appropriate correction factor/term into account – see Clause 9.4.2) is not valid for all sites in the Member State.

The generalisation of equivalence to include other locations than those tested, in which the Equivalent Method is used, is a separate and essential component that should be in the report from the Competent Authority to the Commission (Reference [1] Clause 10). Generally, this report

must show that the equivalence demonstrated at the locations where it was established, applies to all monitoring sites that use this Equivalent Method for which results are being provided to the Commission.

Developing a detailed procedure for generalization of equivalence claims is beyond the scope of this Report. There is no objective procedure for delineating the monitoring sites where a demonstrated equivalence is valid and where it is not. Instead, expert judgement, based on the similarities in conditions that prevail at the various relevant locations, is needed for this.

There are several relevant ways of describing the sites where a demonstrated equivalence is valid. The sites may be classified in similar groups of locations using *station types* (that are characterized primarily by the nearby sources). The validity range of a demonstrated equivalence can also be described by listing the *regions* (parts of the Member State) of validity. A *combination of station types and regions* (e.g. rural stations in regions A, B and C) may also be a useful way. From this description, a *list of stations with the correction factors applied (or equations used)* can be derived and tabled in the report to the Commission (Reference [1] Clause 10).

5.4.3 Extent of tests required

Within this Report, the extent of equivalence testing is specified on the basis of degree of differences between the Standard Method and the Candidate Method.
These differences can – in principle – be separated into two groups (defined subsequently in this Report as 'variations on a theme' and 'different methodologies').

5.4.3.1 *Variations on a theme*

Minor parts of the Standard Method can be modified resulting in 'variations on a theme'.

Examples of 'possible variations':

- The use of different converters to tranform nitrogen dioxide into nitric oxide in chemiluminescence analysers;
- The use of different scrubbers for ozone;
- The use of different sampling media/substrates, e.g., sorbents and filter types;
- The use of different procedures for analyte recovery, e.g., for recovery of benzene from sorbent tubes, and metals and polycyclic aromatic hydrocarbons (PAH) from PM samples;
- The use of calibration procedures differing in the number and/or analyte contents of calibration standards, the type of calibration function and its establishment and the re-calibration frequency;
- The use of different analytical procedures, e.g., modifications to the chromatographic separation for benzene and PAH analysis, and to the atomic spectrometric conditions for metals analysis;
- The use of different PM_{10} filter storage procedures;
- The use of automated filter changers for manual PM_{10} samplers.

5.4.3.2 *Different methodologies*

A Candidate Method may be based on a different measurement principle. Possible examples of different principles are:

- Automated Measurement Systems for benzene using ultraviolet spectrometry as the detection technique;

- Sampling of particulate matter using a sampling inlet with size and shape differing from those specified in PM_{10} and $PM_{2.5}$ norms for the reference sampler;

- Measurement of particulate matter using automated methods, e.g., based on β-ray attenuation or on oscillating microbalances;

- Use of in-situ optical measurement techniques for particulate matter;

- Use of different analytical techniques for the measurement of relevant compounds in sample extracts, e.g., liquid chromatography for benzene, inductively-coupled plasma – optical emission spectrometry for metals;

- Measurement of gases and vapours using diffusive sampling instead of pumped sampling or automated methods;

- Automated measurement of gases based on a different spectrometric technique, e.g., fourier-transform infrared spectrometry (FTIR) for sulphur dioxide;

- Measurement of gases using pumped sampling instead of automated methods.

5.4.3.3 *Practical implications*

In practice, the possible use of different methodologies is limited. Based on practical potential/current applications, the following may be considered as relevant examples:

Sulphur dioxide, nitrogen dioxide, carbon monoxide, ozone
The Reference Method is continuous spectrometry. Candidate Methods of practical value include:
- diffusive sampling with subsequent sample analysis;
- 'flow-through' spectrometric techniques using different principles than the reference technique.

Benzene
The Reference Method is pumped sampling (automated or non-automated) followed by sample analysis using gas chromatography. Candidate Methods of practical value are:
- diffusive sampling with subsequent sample analysis;
- 'flow-through' spectrometry;
- automated measurement using ultraviolet spectrometry after sample enrichment.

EN Standard Methods exist for the measurement of benzene by diffusive sampling and analysis by gas chromatography after thermal or solvent desorption of benzene samples (prEN 14662 parts 4 and 5; refs. 16,17). These methods have been evaluated to be equivalent to the Reference Method within the frame of the validation of all EN Standard Methods for the measurement of benzene (prEN 14662 part 1-5, refs. 13-17).

Particulate matter
The Reference Method is manual pumped sampling onto a filter substrate using a pre-specified aerosol classifier followed by gravimetric analysis. Candidate Methods may be based on:
- use of automated methods, i.e., ß-ray attenuation or (tapered-element) oscillating microbalance;
- use of aerosol samplers with inlet specifications differing from those of the reference sampler.

Metals, benz[a]pyrene
The Reference Method is based on sampling of the PM_{10} aerosol fraction of the total suspended particulate matter in ambient air, with subsequent analysis using atomic absorption spectrometry or inductively-coupled plasma mass spectrometry (metals), or gas or liquid chromatography (benz[a]pyrene). The Candidate Methods may be based on:

- use of alternative analytical techniques;
- use of alternative aerosol samplers (see under particulate matter).

5.5 *Practical approach to equivalence testing*

In principle, the approach to equivalence testing described in this report comprises 4 phases, i.e.:

- An initial non experimental pre-assessment to check whether the Candidate method has the potential for fulfilling the Data quality Objectives in the directives on data capture and measurement uncertainty
- Assessment of the uncertainty of the Candidate Method using an approach based on the principles of ENV 13005 (clause 8) in a series of laboratory tests
- The performance of a series of field tests for confirmation of the findings of the laboratory tests in which the Candidate Method is tested side-by-side to the Standard Method; the 'lack-of-comparability' is tested on the basis of the performance of linear regression with symmetric treatment of both variables, i.e., with uncertainties attributed to both variables
- The evaluation of the resulting uncertainties by comparison of
  - laboratory uncertainty and the uncertainty data quality objective
  - field uncertainty and laboratory uncertainty
  - field uncertainty and the uncertainty data quality objective.

This approach has the advantage that – in the case of 'variations on a theme' – only those contributions to uncertainty that arise from the variation need to be assessed. For example, if a new extraction agent is used, the uncertainty contributions to be tested are the extraction efficiency, blank levels and analytical selectivity. This implies *a priori* knowledge of the uncertainty contributions of all relevant uncertainty sources in the Standard Method. In addition, for manual Candidate Methods for which only the analytical principle but not the sample preparation component differs from the Standard Method (e.g., the use of ICP-OES for the analysis of metals) only the contributions relevant to the use of the different analytical method need to be quantified.

An exception to this is made for the CEN standards covering automated measurement systems for gases; for these, all relevant uncertainty sources must be assessed in order to avoid the use of the equivalence procedure as an route for monitors that have failed the test criteria of the EN standards for automated measurement systems for these species being accepted as equivalent.

In general, for particulate matter the test programmes are restricted to field tests only [8].

It should be noted that measurement procedures based on separate sampling and analysis may be open to 'variations' in parts of the procedure that can lead to systematic differences in measurement results produced by different laboratories on 'identical' air samples. This has been shown to introduce a significant additional contribution to measurement uncertainty –that due to inter-laboratory variability. Consequently, where necessary, the test procedure shall involve more than one laboratory in order to evaluate the contributions to uncertainty from 'between-laboratory' variations.

It should also be noted that application of the approach described in this Report is not mandatory. Other approaches that are in conformity with the requirements of ENV 13005 can also be used, provided that the user can prove the validity of the alternative approach.

5.6 *Operation of the Candidate Method*

The practical operation of the Candidate Method – and possible subsequent Equivalent Method - shall be subject to an appropriate regime of ongoing quality assurance/quality control (QA/QC). This regime shall be documented in the Standard Operating Procedure describing the operation of the method.
Minimum requirements for ongoing QA/QC shall be as reliable as the requirements given in appropriate EN Standard Methods for automated or manual methods [8-18].

In addition, it is recommended that field tests are performed periodically, by operating Reference and Equivalent Methods in parallel, in order to check whether the measurement results remain equivalent.

6 SELECTING A TEST PROGRAMME

The selection and outline of test programmes is given below.

6.1 *General*

Figure 3 gives a flow scheme for selection of the appropriate test programme for any Candidate Method. Four different test programmes have been elaborated for four distinct situations. The distinctions are based in principle on whether:

1. There are 'stated references' that exist for the establishment of measurement traceability, or the extent to which it is possible to quantify all contributions to measurement uncertainty from comparisons starting from primary measurement standards (ENV 13005).
2. The measurement methodology is automated or manual, i.e., based on separate sampling and analysis.

The consequences of these distinctions are explained below.

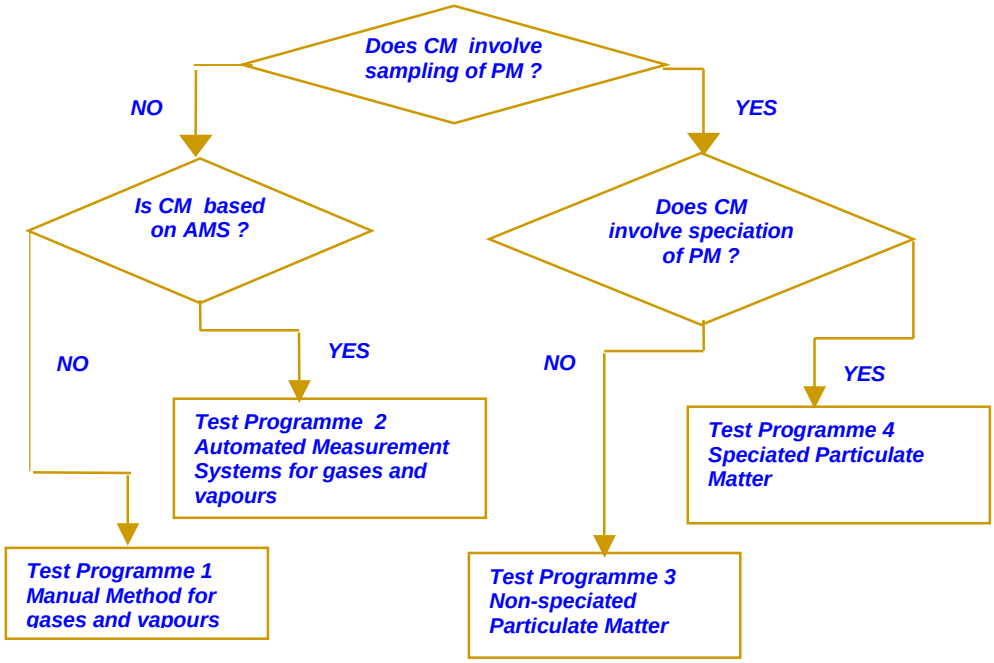


Figure 3. Flow scheme for selection of test programme

6.2 *Measurement methodology*

Test procedures will differ for automated and manual methods for the measurement of gases; for automated methods the method will be tested more or less as a 'black box' (e.g., [9]); for manual methods separate steps in the measurement procedure will be subject to uncertainty evaluation in the laboratory tests (e.g., [13]).

6.3 *Measurement traceability*

The structure and contents of the test programmes given here are determined by the extent to which measurement results can be made traceable to SI units. The existence of primary measurement standards or certified reference materials enable laboratory tests to be performed in which these standards and materials can be used to evaluate measurement bias.

For gaseous and vaporous compounds measurement results can be made fully traceable to SI units through existing primary measurement standards prepared in accordance with ISO 6142, ISO 6144 or ISO 6145. This situation applies to continuous measurements of sulphur dioxide, nitrogen oxides, carbon monoxide and benzene.

For ozone, UV photometry is defined, by convention, as an 'absolute' measurement methodology. A UV photometer of which the measurement uncertainty has been evaluated from first principles may be termed a 'reference' photometer.

For measurements of benzene using pumped sampling methods, reference materials and standards exist through which both the results of the sampling and the analysis can be made fully traceable to SI units.

For heavy metals and benz[a]pyrene reference materials are available which provide traceability for the analytical component of the measurement procedure. However, these generally have sample matrices and measurand concentrations that differ considerably from those relevant to the implementation of the EU Directives. For example, available reference materials for speciated PM measurements – such as NIST SRM 1648 and 1649a – differ in matrix (bulk sample instead of filter), particle size (up to 125 μm) and composition from the reference materials that would be required. Representative reference materials currently do not exist.

For the measurement of particulate matter a more complicated situation exists as no relevant metrological standards or reference materials exist for establishing the traceability of PM_{10} and $PM_{2.5}$ measurements to SI units. Results of measurements of sample volume and sampled mass of particulate matter can be made traceable to SI, but there is no suitable primary standard available to assess the contribution of other uncertainty components of the measurement method. The uncertainty of any Candidate Method therefore has to be determined with reference to a PM reference sampler (e.g., High-Volume Samplers or Low-Volume Samplers as specified in EN 12341 for PM_{10}), assuming these 'reference samplers' to be unbiased with respect to the applied particle-size convention.

6.4 *Specification of test programmes*

Test Programme 1 refers to manual methods for gases and vapours (benzene, carbon monoxide, sulphur dioxide, nitrogen dioxide and ozone).

- Test Programme 1A: Laboratory test programme for variations on the Standard Method; laboratory and field test programme for pumped sampling alternatives to standard methods for other gaseous pollutants
- Test Programme 1B: Laboratory and field test programmes for diffusive sampling analogous to test programmes of EN 13528.

Test Programme 2 refers to alternative automated measurement systems for gases and vapours, (benzene, carbon monoxide, sulphur dioxide, nitrogen dioxide and ozone) e.g., using other spectrometric techniques.

Test Programme 3 refers to alternative methodologies for the monitoring of non-speciated particulate matter. Test programme 3 includes testing of a size selective inlet, when this differs from that of the PM reference sampler.

Test Programme 4 refers to the determination of speciated particulate matter (metals and benz[a]pyrene in samples of particulates).

7 TEST PROGRAMME 1 - MANUAL METHODS FOR GASES AND VAPOURS

7.1 General

This test programme describes a procedure for determining whether a Candidate Method (CM) is suitable to be considered equivalent to the EN Standard Method based on one of the Reference Methods for the measurement of gases and vapours in ambient air [9-15], using manual measurement methods (with separate sampling and analysis). These EN Standard Methods have been developed to meet the Data Quality Objectives of the Air Quality Framework Directive and its relevant Daughter Directives. Therefore CMs will similarly have to meet these Directives' requirements.

This test programme is suitable for evaluating:

- pumped and diffusive sampling methods as alternatives for automated methods for the measurement of sulphur dioxide, nitrogen dioxide, carbon monoxide, ozone and benzene
- diffusive sampling methods and modified pumped sampling methods as alternatives for benzene.

7.2 Overview of the test procedures

The EN Standard Methods based on the Reference Methods [9-15] specify procedures to determine the expanded uncertainty of the method from the determination of uncertainty components obtained during the specific tests. This expanded uncertainty is compared with the expanded uncertainty criterion given in the relevant Daughter Directive.

Testing for equivalence will normally be carried out in two parts: a laboratory test in which the contributions of the different uncertainty sources to the measurement uncertainty will be assessed, and a field test in which the Candidate Method will be tested side-by-side with the relevant Standard Method.

If a CM is a modification to an existing EN Standard Method, then only the laboratory performance characteristics that are affected by the modification need to be tested and their standard uncertainties calculated. The standard uncertainties associated with the affected performance characteristics shall then be used together with these existing standard uncertainties for the other characteristics, to determine again the combined measurement uncertainty, $u_{c.}$

If a CM utilises a measurement method that is different to a Standard Method, then all of the tests shall be performed.

In both cases the results of existing studies, when demonstrably obtained according to the requirements of this test procedure, may be used to determine standard uncertainties.

The CM should be tested in a way that is representative of its practical use; for example, the frequencies of tests (e.g., response drift) and re-calibrations (e.g., flow rates) that are used in practice should be applied in the test programmes.

For diffusive sampling methods for benzene, information on uncertainty sources exists in EN norms [16,17]; these norms should be consulted when alternative diffusive sampling methods are considered as Candidate Methods. For diffusive sampling of inorganic gases, no such information is currently available in this form. It is necessary to compile and evaluate this information in the course of the validation of diffusive sampling methods for these gases.

Test Programme 1 consists of a laboratory and field test programme. The Laboratory Test Programme is separated into two parts (1A and 1B), covering methods for which the volume of air sampled can be made traceable to SI units (pumped sampling) and to methods for which this is not possible (diffusive sampling).

Candidate Methods must pass the criteria for the Laboratory Test programme, and also pass the criteria for the Field Test programme. Only Candidate Methods that pass the Laboratory Test Programme shall proceed to the Field Test Programme.

7.3 Laboratory Test Programme

In the laboratory test programme, the uncertainty sources listed in Table 7.1 are considered and assessed, where appropriate.

Table 7.1. Laboratory Test Programme 1: Uncertainty sources

Uncertainty source	**Symbol**	
	Pumped sampling	***Diffusive sampling***
1 *Sample volume*	V_{sam}	
1.2 Sample flow / uptake rate 1.2.1 calibration and measurement 1.2.2 variation during sampling	φ	υ
1.3 Sampling time	t	t
1.4 Conversion to standard temperature and pressure		
2 *Mass of compound in sample*	m_{sam}	m_{sam}
2.1 Sampling efficiency	E	*
2.2 Compound stability	A	A
2.3 Extraction/desorption efficiency	D	D
2.4 Mass of compound in calibration standards	m_{CS}	m_{CS}
2.5 Response factors 2.5.1 lack-of-fit of calibration function 2.5.2 analytical repeatability 2.5.3 drift between calibrations	 F d	 F d
2.6 Selectivity	R	R
3 *Mass of compound in blank*	m_{bl}	m_{bl}

** For diffusive sampling, sampling efficiency will be incorporated in the uptake rate.*

The uncertainty sources that require assessment depend on the differences between Candidate and Standard Methods as follows:

Is the Candidate Method based on a different measurement principle?
In that case, the full Test Programme needs to be performed.

Does the sampling principle of the Candidate Method differ from that of the Standard Method (e.g. diffusive instead of pumped sampling for benzene)?
In this case, uncertainty source 1.2 needs to be assessed.

Does the analytical principle of the Candidate Method differ from that of the Standard Method, with the sampling being the same?
In this case, the uncertainty sources under 2.5, 2.6 and 3 need to be assessed.

Is the Candidate Method a modification of the Standard Method?
In this case, the uncertainty sources relevant to the modification need to be investigated, e.g.

- 2.1, 2.2, 2.3 and 3 for alternative sorbents

- 2.3 and 2.6 for alternative extraction solvents
- 2.5 and 2.6 for alternative analytical configurations.

7.3.1 Test programme 1A: pumped sampling

7.3.1.1 *Sampled volume of air*

The sampled volume of air shall be sufficient to allow reliable quantification of the pollutant concentration at the lower end of the measurement range (10% of the limit value).
In practice, the sampled volume of air may be determined in two ways:

- on the basis of a sample flow rate measured before sampling as

$$V_{sam} = \varphi_{start}\, t \qquad (7.1a)$$

- on the basis of measuring the sample flow rate directly before and after sampling as

$$V_{sam} = \frac{(\varphi_{start} + \varphi_{end})}{2}\, t \qquad (7.1b)$$

in which
φ_{start} = sample flow rate before sampling, calculated as the average of ≥ 3 consecutive measurements
φ_{end} = sample flow rate after sampling, calculated as the average of ≥ 3 consecutive measurements
t = sampling time.

The first situation will occur in monitoring networks in which sequential samplers are used that are only checked or re-calibrated after prolonged intervals (e.g. 6 months). These samplers mostly use mass-flow controllers.

The uncertainty in the volume of air sampled is made up of contributions from
- the measurements of the flow rates before, or before and after, sampling
- the measurement of the sampling time
- flow rate drift, or variations in the flow rate during the sampling period.

For the two cases given in eq. (7.1a) and (7.1b) the uncertainty of the sampled volume u(V) may be derived:

$$\frac{u^2(V_{sam})}{V_{sam}^2} = \frac{u^2(\varphi_{start})}{\varphi_{start}^2} + \frac{u^2(t)}{t^2} + \frac{\Delta^2\varphi}{3\varphi_{start}^2} \qquad (7.2a)$$

$$\frac{u^2(V_{sam})}{V_{sam}^2} = \frac{u^2(\varphi_{start}) + u^2(\varphi_{end})}{(\varphi_{start} + \varphi_{end})^2} + \frac{u^2(t)}{t^2} + \frac{\Delta^2\varphi}{\left[12\,\frac{(\varphi_{start} + \varphi_{end})}{2}\right]^2} \qquad (7.2b)$$

where
$\Delta\varphi$ = flow rate drift. i.e. the difference between two flow subsequent rate measurements:

$$\Delta\varphi = \varphi_{start} - \varphi_{end} \qquad (7.3)$$

$u(\varphi_{start})$ is the standard uncertainty in the measurement of the flow before sampling (see 7.3.1.1.1)

$u(\varphi_{end})$ is the standard uncertainty in the measurement of the flow after sampling (see 7.3.1.1.1)

u(t) is the standard uncertainty in the measurement of the time (see 7.3.1.1.2)

In the situation where only the flow rate before sampling is measured, the drift in flow rate over the period of unattended operation should have been established in a test programme preceding the practical use of the sampler.

Because conversion to standard temperature and pressure (STP) is prescribed in the relevant 'Daughter Directives', uncertainty contributions for this conversion are to be taken into account. These contributions will depend on whether mass-flow controlled or volume-controlled sampling devices are used. The calculation of individual uncertainty contributions is given in 7.3.1.1.3.

7.3.1.1.1 Sample flow calibration and measurement

The uncertainty in the measurement of the flow rates before and after sampling is calculated from the uncertainty in the readings of the flow meter used which can be derived from calibration certificates, assuming the calibration is fully traceable to primary standards of flow, and the uncertainty of the actual flow rate measurement results, as

$$\frac{u^2(\varphi)}{\varphi^2} = \frac{u_{cal}^2 + \frac{s_{meas}^2}{n}}{\varphi^2} \qquad (7.4)$$

where
$u(\varphi)$ is the standard uncertainty in the measurement of flow
u_{cal} = uncertainty due to calibration of the flow meter
s_{meas} = standard deviation of individual flow measurements, determined from ≥ 3 measurements
n = number of flow measurements performed under practical conditions of use.

7.3.1.1.2 Sampling time

The sampling time t should be measured to within ± 0,5 min. Then for a sampling time of 8 hours or more the relative uncertainty due to the measurement of *t* is negligible.

7.3.1.1.3 Conversion of sample volume to STP

Mass-flow controlled sampling devices
For mass-controlled sampling devices a conversion of the sample volume to STP may be affected by direct conversion of measured flow rates to values at STP. For conversion, the following equation is used:

$$\varphi_{STP} = \varphi \frac{P}{101{,}3} \frac{293}{(T + 273)} \qquad (7.6)$$

where
φ_{STP} = sample flow converted to STP
φ = actual measured sample flow
P = actual air pressure during the flow measurements (in kPa)
T = actual air temperature during the flow measurements (in °C).

By modification of Eq. (7.1) through substitution of ϕ with ϕ_{STP} , the sample volume converted to STP is:

$$V_{sam,STP} = \varphi_{start,STP} t \qquad (7.7a)$$

$$V_{sam,STP} = \frac{\left(\varphi_{start,STP} + \varphi_{end,STP}\right)}{2} \cdot t \qquad (7.7b)$$

The uncertainty contribution for mass-flow controlled sampling devices can then be obtained by extending equation (7.4) to:

$$\frac{u^2(\varphi_{STP})}{\varphi_{STP}^2} = \frac{u_{cal}^2 + \frac{s_{meas}^2}{n}}{\varphi^2} + \frac{u^2(P)}{P^2} + \frac{u^2(T)}{T^2} \qquad (7.8)$$

where
φ_{STP} = sample flow corrected to STP
$u(\varphi_{STP})$ = uncertainty in the sample flow corrected to STP
u_{cal} = uncertainty due to calibration of the flow meter
s_{meas} = standard deviation of individual flow measurements, determined from a minimum of 3 measurements
n = number of flow measurements performed under practical conditions of application
u(T) = uncertainty of the actual air temperature value during the flow measurements
u(P) = uncertainty of the actual air pressure value during the flow measurements
P = actual air pressure during the flow measurements
T = actual absolute air temperature during the flow measurements.

By substitution of φ and $u(\varphi)$ by φ_{STP} and $u(\varphi_{STP})$, respectively, in Eq. (7.2), the uncertainty of the sample volume, converted to STP, when employing mass-flow controlled sampling devices is obtained directly as:

$$\frac{u^2\left(V_{sam,STP}\right)}{V_{sam,STP}^2} = \frac{u^2\left(\varphi_{start,STP}\right)}{\varphi_{start,STP}^2} + \frac{u^2(t)}{t^2} + \frac{\Delta^2\varphi_{STP}}{3\varphi_{start,STP}^2} \qquad (7.9a)$$

$$\frac{u^2\left(V_{sam,STP}\right)}{V_{sam,STP}^2} = \frac{u^2\left(\varphi_{start,STP}\right) + u^2\left(\varphi_{end,STP}\right)}{\left(\varphi_{start,STP} + \varphi_{end,STP}\right)^2} + \frac{u^2(t)}{t^2} + \frac{\Delta^2\varphi_{STP}}{\left[12\,\frac{\left(\varphi_{start,STP} + \varphi_{end,STP}\right)}{2}\right]^2} \qquad (7.9b)$$

Volume-controlled sampling devices
When using volume-flow controlled sampling devices, knowledge is required of the mean ambient temperature and pressure that occurs during sampling. These are used as follows for the conversion:

$$V_{sam,STP} = V_{sam}\,\frac{\overline{P}}{101{,}3}\,\frac{293}{\left(\overline{T} + 273\right)} \qquad (7.10)$$

where
$\overline{P}$ = average air pressure during the sampling period (in kPa)
$\overline{T}$ = average air temperature during the sampling (in °C).

Uncertainties in values of $\overline{T}$ and $\overline{P}$ used for conversion may be obtained from

- actual measurements, taking into account the uncertainty in the temperature and pressure measurements
- knowledge of extremes of temperature and pressure during sampling, assuming these to be uniformly distributed.

For example, if the temperature extremes are known to be T_{min} and T_{max} the uncertainty in $\overline{T}$ may be calculated from

$$u^2(\overline{T}) = u_{cal}^2 + \frac{(T_{max} - T_{min})^2}{12} \quad (7.11)$$

where
u_{cal} = uncertainty due to calibration of the temperature meter.

Generally, the first term will be negligible compared to the second.

The above uncertainty contributions are then combined to give the uncertainty in the sample volume converted to STP for volume-controlled sampling devices as:

$$\frac{u^2(V_{sam,STP})}{V_{sam,STP}^2} = \frac{u^2(V_{sam})}{V_{sam}^2} + \frac{u^2(\overline{T})}{\overline{T}^2} + \frac{u^2(\overline{P})}{\overline{P}^2} \quad (7.12)$$

7.3.1.2 *Mass of compound sampled*

The mass of a compound sampled may be expressed as:

$$m_{sam} = \frac{m_{meas}}{E \cdot A \cdot D} \quad (7.13)$$

where
E = sampling efficiency
A = compound stability in the sample
D = extraction/desorption efficiency
m_{meas} = measured mass of compound in the analytical sample (extract, desorbate) before correction.

A correction for extraction/desorption efficiency shall be applied when D is significantly different from 1 (see 7.3.2.1.3).

7.3.1.2.1 Sampling efficiency

For the sampling medium to be used the breakthrough volume shall be determined under reasonable worst-case conditions. In practice, these conditions will consist of a combination of a high concentration, high temperature, high air humidity, and the presence of high levels of potentially interfering compounds. As the worst-case conditions will vary between sample locations, test conditions may be adapted to these local conditions.

The sample volume shall be less than half the experimentally established breakthrough volume.

In that case the sampling efficiency will be 100% and will not contribute to the uncertainty in m_{sam}.

7.3.1.2.2 Compound stability

The compound stability shall be established experimentally through storage under conditions (time, temperature, environment) that are typical to the individual monitoring network. Tests shall be performed at a compound level corresponding to the ambient air limit or target value.

At times t=0 and t=t, n samples shall each be analyzed under repeatability conditions ($n \geq 6$). For both times the samples shall be randomly selected from a batch of representative samples in order to minimize possible systematic concentration differences. As a test of (in)stability, a t-test will be performed (95% confidence, 2-sided). The t-test must show no significant difference between results obtained at the start and end of the stability test.

The uncertainty of the stability determination consists of contributions from:

- extraction/desorption (random part of extraction/desorption efficiency)
- calibration (random part of calibration)
- analytical precision
- inhomogeneity of the sample batch.

However, the uncertainty contribution of the determination of stability will already be covered by contributions determined in Clause 7.3.1.2.3 and it therefore does not need to be taken into account separately.

7.3.1.2.3 Extraction/desorption efficiency

The extraction/desorption efficiency of the compound from the sample and its uncertainty are typically obtained from replicate measurements on Certified Reference Materials (CRMs). The uncertainty due to incomplete extraction/desorption for the level corresponding to the limit value is calculated from contributions of

- the uncertainty in the concentration of the CRM
- the standard deviation of the mean mass determined

as

$$\frac{u^2(D)}{D^2} = \frac{u^2(m_{CRM}) + \frac{s^2(m_D)}{n}}{m_{CRM}^2} \qquad (7.14)$$

where
m_{CRM} = certified mass in the CRM
$s(m_D)$ = standard deviation of the replicate measurement results of the mass determined
n = the number of replicate measurements of the CRM.

When D is significantly different from 1 (at the 95% confidence level), the measurement result shall be corrected accordingly (see eq. (7.1)).

The value of $s(m_D)$ is used as an indicator of the relative uncertainty due to analytical repeatability w_{anal}:

$$w_{anal}^2 = \frac{s^2(m_D)}{m_D^2} \qquad (7.15)$$

7.3.1.2.4 Corrections to the measured mass of the compound

The uncertainty in the measured mass of a compound is determined by

- the uncertainty in the concentrations of the calibration standards used
- the lack-of-fit of the calibration function

- drift of detector response between calibrations
- the precision of the analysis
- the selectivity of the analytical system used.

Calibration standards

The uncertainty of the concentration of a compound in the calibration standards used will depend on the type of calibration standard used. For a tube standard prepared by sampling from a standard atmosphere it will depend on:

- the uncertainty of the concentration in the generated standard atmosphere; uncertainty assessments for this parameter can be found in ISO 6144 and 6145 [19,20]
- the uncertainty of the sampled volume of the standard atmosphere.

The uncertainty is calculated as

$$\frac{u^2(m_{cs})}{m_{cs}^2} = \frac{u^2(C_{sa})}{C_{sa}^2} + \frac{u^2(V)}{V^2} \qquad (7.16)$$

where
$u(m_{cs})$ = uncertainty in the mass in the calibration standard (m_{cs})
$u(C_{sa})$ = uncertainty in the concentration in the standard atmosphere (C_{sa})
u(V) = uncertainty in the volume of the standard atmosphere sampled (V).

For calibration standards consisting of solutions the uncertainty will be built up of contributions from:

- the purity of the compound used as calibrant; as the compounds under study are generally available in purities > 99%, the contribution of the purity may be considered insignificant
- when gravimetry is used to prepare the calibration solutions: the uncertainties in the weighings of compounds and solutions
- when volumetric techniques are used to prepare the calibration solutions: the uncertainties in the calibrated volumes of glassware and syringes used.

Note: Examples of calculations of uncertainties can be found in refs. [21] and [22].

For tube standards prepared by spiking from a solution and subsequent purging of the solvent, the uncertainty is composed of the uncertainties of the compound concentration in the solution, the spiking volume, the sampling efficiency and possible selectivity effects due to the presence of residual solvent.

Lack-of-fit of calibration function

The relative uncertainty due to lack-of-fit of the calibration function can be calculated for the relevant concentration (corresponding to the mass of benzene sampled at the limit value) from parameters obtained by a least-squares linear regression ($r = a + b.m_{cs}$), weighted in the concentration of the calibration standard.

Note: Options for the calculation of the uncertainty are given in ref. [21].

As a worst-case approach, the relative uncertainty shall be estimated as

$$w_F^2 = \frac{u^2(m_r)}{m_r^2} = \frac{u^2(r) + s^2(a) + s^2(b)m_r^2}{b^2 m_r^2} \qquad (7.17)$$

where
m_r = mass calculated from the regression equation at response r

u(r) = the uncertainty of the response r
b = slope of calibration function
a = intercept of calibration function
s = standard deviation of parameter between parentheses.

Response drift between calibrations

Normally, the current response factor will be used until a new one is established. In the interval between the re-establishment of its uncertainty, response checks – and, when necessary, adjustments of response factors - shall be performed as an element of ongoing quality control.

In the interval before the next checks response drift may occur. The relative uncertainty due to response drift for the period between subsequent adjustments of response factors shall then be estimated from data on the relative differences in responses between subsequent checks, as

$$w_d^2 = \frac{(r_n - r_{n-1})^2}{3\left(\frac{r_n + r_{n-1}}{2}\right)^2} \qquad (7.18)$$

where r_n is the detector response for a calibration standard corresponding closest to the mass representing a sample at the limit value. This approach assumes that no correction is applied for response drift, e.g., by averaging of subsequently determined response factors.

Selectivity

The analytical system used shall be optimized in order to minimize uncertainty due to the presence of potential interferents. Tests shall be performed with typical interferents at levels corresponding to 5 times the limit value of the compound under study. The uncertainty due to interferences may be obtained from ISO 14956 [23] as

$$w_R^2 = \frac{(r_+ - r_0)^2}{3r_0^2} \qquad (7.19)$$

where r_+ represents the response with interferent, and r_0 represents the response without.

7.3.1.2.5 Combined uncertainty in the sampled mass

The contributions given above are combined to give the uncertainty of the mass of compound in the air sample as

$$\frac{u^2(m_{sam})}{m_{sam}^2} = \frac{u^2(m_{cs})}{nm_{cs}^2} + w_{anal}^2 + w_F^2 + w_d^2 + w_R^2 \qquad (7.20)$$

where
n = number of calibration standards used to construct the calibration function (≥5)
w_R = relative uncertainty due to (lack of) selectivity of the analytical system.

7.3.1.3 Mass of compound in sample blank

The mass of compound in a sample blank is determined by analysis under repeatability conditions of a series of sample blanks; a minimum of 6 replicate analyses should be performed. The uncertainty is then calculated using the slope of the calibration function extrapolated to the blank response level as

$$u^2(m_{bl}) = \frac{s_{bl}^2}{nb_{bl}} \quad (7.21)$$

where
s_{bl} = standard deviation of the replicate blank analyses
n = number of replicate analyses
b_{bl} = slope of the calibration function at the blank response level.

When the blank response is less than 3 times the noise level of the detector, then the blank level and its uncertainty shall be calculated from the detector noise level using the slope of the calibration function extrapolated to zero response assuming a uniform distribution, as

$$m_{bl} = \frac{3r_0}{2b_0} \quad (7.22)$$

$$u^2(m_{bl}) = \frac{9r_0^2}{12} \quad (7.23)$$

where
r_0 = noise level
b_0 = slope of calibration function at zero response.

7.3.1.4 Combined uncertainty

The combined relative uncertainty of the compound concentration in the air sampled is obtained by combination of contributions given in Clauses 7.3.1.1-7.3.1.3 as

$$w_{CM,lab}^2 = \frac{u_c^2(C_m)}{C_m^2} = \frac{u^2(V_{sam,SPT})}{V_{sam,SPT}^2} + \frac{u^2(m_{sam}) + u^2(m_{bl})}{(m_{sam} - m_{bl})^2} \quad (7.24)$$

7.3.1.5 Expanded uncertainty

The expanded relative uncertainty of the Candidate Method resulting from the laboratory experiments, $W_{CM,lab}$ at the 95% confidence level is obtained by multiplying $w_{CM,lab}$ with a coverage factor appropriate to the number of degrees of freedom of the dominant components of the uncertainty resulting from the performance of the test programme. This can be calculated by applying the Welch-Satterswaithe equation (ISO-GUM, H2). For a large number of degrees of freedom, a coverage factor of 2 is used.

Note: as a first approximation, the number of degrees of freedom may be based on that of an uncertainty contribution covering more than 50% of the variance budget.

7.3.1.6 Evaluation of results of the laboratory tests

The resulting $W_{CM,lab}$ is compared with the expanded relative uncertainty based on the data quality objective for the relevant species W_{dqo}.
If $W_{CM,lab} \leq W_{dqo}$, the field test programme can be performed; if not, the Candidate Method shall first be improved, and relevant changes tested in the laboratory test programme.

7.3.2 Test Programme 1B. Diffusive sampling

7.3.2.1 Reduced test programme

For general information about testing of diffusive samplers, the reader is referred to EN Norms EN 13528 parts 1-3 [24-26].

As a first estimate, the diffusive sampling flow (uptake rate) υ and its uncertainty can be determined under 2 sets of extreme conditions [27]. Extreme conditions for diffusive sampling are characterized by extremes of

- Temperature (low and high): these will depend on prevailing local or regional conditions and will differ between member states. Member states must cover prevailing regional extremes.
- Relative humidity (as for temperature)
- Air velocity: this should always be within the range required for proper functioning of the sampler. This range shall be established beforehand; in practice, adherence to the maximum velocity shall be ensured through use of appropriate wind shields (see EN 13528 part 3, [25]). In the tests, a default level of approximately 0,5 m s^{-1} is applied
- Concentrations of interferents: interferents will either affect the concentration of the compound of interest or compete for sorption sites with the compound of interest. Interferents and maximum extremes will depend on prevailing local or regional conditions. Member states must cover prevailing regional extremes in their test programmes.

Test conditions consist of:

- Extreme 1: A combination of high temperature, high relative humidity and high concentrations of interferents
- Extreme 2: A combination of low temperature, low relative humidity and low concentrations of interferents.

In each test, a minimum of 6 samplers is exposed for the exposure period considered.

The resulting characteristics to be derived are υ_{high} , s_{high} , υ_{low} and s_{low}.

The effective sampling (uptake) rates and their uncertainties are calculated as follows:

$$\upsilon_{eff} = \frac{\upsilon_{high} + \upsilon_{low}}{2} \tag{7.25}$$

$$u^2(\upsilon_{eff}) = \frac{\left[\left(\upsilon_{high} - \upsilon_{low}\right) + 2\frac{s_{high}}{\sqrt{n_{high}}} + 2\frac{s_{low}}{\sqrt{n_{low}}}\right]^2}{24} \tag{7.26}$$

where
s_{high} is the standard deviation of the determination of the uptake rate under conditions Extreme 1
s_{low} is the standard deviation of the determination of the uptake rate under conditions Extreme 2
n represents the number of samplers exposed in each situation.

The uncertainty calculated in this way is based on the assumption of a triangular distribution of values of υ and provides a 'first' uncertainty estimate. The uncertainty assessment can be refined – if necessary – through the performance of extended tests.

7.3.2.2 *Extended test programme*

In the extended test programme, the factors affecting the sampling rate (see above) are varied in 2-level (high/low) or 3-level (high/medium/low) experimental designs. The number of experiments to be performed can be based on an orthogonal or 'Taguchi' design. For the 3-factor/2-level

approach a minimum number of 4 experiments suffices, for a 3-factor/3-level design 9 experiments are needed.

The resulting average sampling (uptake) rate and its uncertainty can be evaluated by applying analysis of variance.

7.4 Field test programme

7.4.1 General

Field tests shall be performed in which the Candidate and a Standard Method based on the Reference Method are compared side-by-side. The measurements will serve to assess

- 'between-sampler' uncertainty of the Candidate Method through the use of replicate samplers
- 'comparability' of the Candidate and Standard Methods.

Generally, results of existing studies, when demonstrably obtained according to the requirements of this test procedure, may be used to determine standard uncertainties. This is particulary relevant to the estimation of between-sampler/instrument uncertainties.
In order to assure proper functioning of the Standard Method, two reference samplers or instruments may be used. In this case the mean squared difference of the results of both reference samplers/instruments can be used as an estimate of the (random) uncertainty of the reference method.

The number of replicate samplers needed to determine the between-sampler uncertainty of the Candidate Method will depend on

- the averaging period of the measurement
- the practicability of performing multiple measurements in parallel
- whether the analysis part of the Candidate Method is to be carried out by more than one laboratory.

Each laboratory carrying out analysis for the Candidate Method shall provide at least two samplers.

When more than one laboratory will carry out the analysis, the field tests shall also be used to evaluate between-laboratory contributions to the uncertainty of the measurement results.

7.4.2 Experimental conditions

Test sites shall be representative of typical conditions for which equivalence will be claimed, including possible episodes of high concentrations. A minimum of 4 comparisons shall be performed with particular emphasis on the following variables, if appropriate:

- Composition of the air, notably high and low concentrations of the measured compound and potential interferents
- Air humidity and temperature (high and low) to cover any effects on sampling efficiency or desorption efficiency
- Wind speed (high and low) to cover any dependency of sampler performance due to deviations from ideal behaviour.

A minimum of 40 measurement results for the Candidate Method per comparison shall be collected over a minimum of 8 sampling periods covering a minimum of 20 days (e.g. 2 samplers and 20 periods, 5 samplers and 8 periods).

Samplers and instruments shall be positioned in such a way that the effect of spatial inhomogeneities in the compound concentration in the sampled air are negligible in comparison with other uncertainty contributions.

Both methods shall be operated under conditions reflecting practical application in the field, e.g., calibration intervals, flow checks, analysis of blank samples.

During the tests, the following information shall be collected and recorded
- Calibration procedures, equipment and intervals
- (Results of) quality checks
- Temperature and pressure of the sampled air
- Other conditions relevant for the measurements performed (e.g., air humidity)
- Particular events/situations that may be of influence on measurement results.

7.4.3 Evaluation of the field test data

7.4.3.1 *Conversion of measurement results to STP*

For the measurement of gaseous pollutants under Directives 1999/30/EC, 2000/69/EC and 2002/3/EC a conversion is required of measurement results to conditions of standard and pressure (STP, 20 °C, 101,3 kPa). Clause 7.3.1.1.3 describes the conversion and the assessment of the resulting uncertainty contribution.

7.4.3.2 *Suitability of the dataset*

Of the full dataset at least 20% of the results shall be greater than or equal to the upper assessment threshold specified in the relevant Daughter Directive.
Data shall only be removed from the data set when sound technical reasons can be found for doing so. All valid data shall be used for further evaluation.

Note: Indications of outlying data within replicate sets may be obtained using Grubb's tests on the individual single-period variances. Tests are to be performed at the 99% level.

7.4.3.3 Calculation of performance characteristics

7.4.3.3.1 Between-sampler/instrument uncertainty

If the Standard Method is based on an AMS, the results for each measurement period i are averaged first to give 24-hour values y_i.

The relative between-sampler uncertainty for individual laboratories w_{bs} is calculated from the differences of results of the candidate samplers/instruments operated in parallel as:

$$w_{bs}^2 = \frac{\sum_{i=1}^{n} \left(y_{i,1} - y_{i,2}\right)^2}{2n\bar{y}^2} \quad \text{for duplicate samplers} \qquad (7.27)$$

where
$y_{i,1}$ and $y_{i,2}$ are the results of parallel measurements for a single period *i*
$\bar{y}$ = average of all measurement results of the Candidate Method
n = number of measurement results.

$$w_{bs}^2 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{p} \left(y_{i,j} - \overline{y_i}\right)^2}{n(p-1)\bar{y}^2} \quad \text{for replicate samplers with p > 2} \qquad (7.28)$$

where
y_{ij} = result of measurement *j* for a single period *i*
$\bar{y}_i$ = mean result for period i
p = number of replicates for period i.

Where more than one analytical laboratory is participating, equation 7.28 shall be used to calculate the between-laboratory w_{bs}.

The w_{bs} between sampler uncertainty component for each individual laboratory and the between-laboratory w_{bs} (if relevant) shall comply with the criteria given in Annex A.

If the performance of a single laboratory causes a method implemented by more than two laboratories to fail the criteria, then the results for this laboratory may be excluded, if sound technical grounds exist for doing so.

7.4.3.3.2 *Comparison with Standard Method*

First, the performance of the reference samplers/instruments is checked by calculation of the relative between-sampler/instrument uncertainty as in eq. (7.27) or (7.28). This relative uncertainty shall be ≤ 3%.

For the comparison of the Candidate Method with the Standard Method, first the results of replicate measurements are averaged to give data pairs 'Candidate Method – Standard Method' with equal measurement periods.

For the evaluation of the uncertainty due to the 'lack of comparability' between Candidate and Standard Method it is assumed that the relationship between measurement results of both methods can be described by a linear relation of the form

$$y_i = a + bx_i \tag{7.29}$$

where x_i is the average result of the Standard Method over period *i.*

The relation between the average results of the Candidate Method and those of the Standard Method is established using a regression technique that leads to a symmetrical treatment of both variables. A commonly applied technique is orthogonal regression [7,28].

The uncertainty due to lack of comparability will be a function of the concentration of the measurand.
The general relationship describing the dependence of $u_{C\text{-}S}$ on x_i is given by

$$u^2_{C\_S}(y_i) = \frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)x_i]^2 \tag{7.30}$$

where
RSS = the sum of (relative) residuals resulting from the orthogonal regression
$u(x_i)$ = random uncertainty of the Standard Method; as such, the value of u_{bs} calculated for the application of the Standard Method in these tests may be used.
Note: The above equation is used in all the equivalence test programmes in this Report for the field evaluation of a CM against the Standard Method. It can be seen that it corrects for the random uncertainty component arising from the implementation of the Standard Method, and thus if this is implemented badly, in parallel with the operation of a high performance CM, small or negative results may be obtained for the uncertainty of the CM using this equation.

Algorithms for the calculation of *a* and *b* and their variances are given in Annex B.

RSS, the sum of (relative) residuals is calculated using eq. 7.31a or 7.32b, depending on whether the residuals or relative residuals are constant.

$$RSS = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \text{ when } (y_i - a - bx_i)^2 \text{ is constant} \quad (7.31a)$$

$$RSS = (a + bx_i)^2 \sum_{i=1}^{n} \left(\frac{y_i}{a + bx_i} - 1 \right)^2 \text{ when } \left(\frac{y_i}{a + bx_i} - 1 \right)^2 \text{ is constant} \quad (7.31b)$$

7.4.3.4 *Calculation of the combined uncertainty of Candidate Method*

The combined relative uncertainty of the Candidate Method $w_{c,CM}$ is calculated by combining the contributions found in 7.4.3.2.1 and 7.4.3.2.2 as follows:

$$w_{c,CM}^2(y_i) = \frac{u_{C\_S}^2(y_i)}{y_i^2} \quad (7.32)$$

In this way, $w_{c,CM}$ is expressed as a function of the compound concentration.

The uncertainty at the limit value w_{CM} is calculated by taking as y_i the concentration at the limit value.

7.4.3.5 *Calculation of the expanded uncertainty of Candidate Method*

The expanded relative uncertainty of the results of the Candidate Method is calculated by multiplying $w_{c,CM}$ by a coverage factor k reflecting the appropriate number of degrees of freedom resulting from the determination of $w_{c,CM}$ as

$$W_{CM,field} = k \cdot w_{c,CM} \quad (7.33)$$

In view of the large number of experimental results available, a coverage factor k=2 can be used for a 95% confidence level.

7.4.4 Evaluation of results of field tests

The resulting uncertainty estimate W_{CM} is compared with the expanded relative uncertainty obtained from the laboratory test programme WCM,l and the expanded relative uncertainty based on the data quality objective for the Standard Method W_{dqo}.
In principle, three cases are possible

1. $W_{CM,field} \leq W_{CM,lab}$: the Candidate Method is accepted as equivalent to the Standard Method

2. $W_{CM,lab} < W_{CM,field} \leq W_{dqo}$: the Candidate Method is accepted conditionally; before final acceptance, the uncertainty evaluation from the laboratory tests should be revisited and corrected such that situation 1 occurs

3. $W_{CM,field} > W_{dqo}$: the Candidate Method is not accepted as Equivalent Method

8 TEST PROGRAMME 2: AUTOMATED MEASUREMENT SYSTEMS FOR GASES

8.1 *General*

This test programme describes a procedure for determining whether a Candidate Method is suitable to be considered equivalent to the EN Standard Method for the measurement of gases and vapours in ambient air, using automated measurement systems. The EN Standard Methods have been developed to meet the Data Quality Objectives of the Air Quality Framework Directive and its Daughter Directives. Therefore, Candidate Methods will similarly have to meet the directives' requirements.

This test programme covers the requirements for the equivalence testing of an AMS where it is practical to achieve measurements that have full traceability to SI units. These include continuous ambient-air analysers monitoring sulphur dioxide, the nitrogen oxides (NO and NO_2) ozone and carbon monoxide. Analysers measuring benzene in ambient air by sequential automated (quasi–continuous) sampling and subsequent measurements by gas chromatography, are also covered.

The use of similar automated methods for the measurement of precursor 'non-methane hydrocarbons' discussed in EU Directive 2002/03/EC are not included because these measurements are not covered by a EN Standard Method, nor are there yet any assigned uncertainty requirements for such methods.

The methodology specified in this Section for equivalence testing follows very closely the procedures specified in the appropriate EN Norms prepared by CEN Technical Committee 264 'Air Quality'.

8.2 *Overview of the test procedures*

The EN Standard Methods specify procedures for the type-approval testing of analysers to determine whether their performance (overall measurement uncertainty, data capture etc) conforms with the requirements of the relevant Daughter Directive. These tests define all the individual performance characteristics which contribute to the combined uncertainty of the method (repeatability, responses to cross interferents etc), and which therefore shall be tested.

The Standard Methods also specify minimum (performance) criteria to which the individual performance characteristics shall conform. In addition, the Standards specify procedures to determine the expanded uncertainty of the method from the component performance criteria obtained during tests, and this expanded uncertainty shall be compared with the expanded uncertainty Data Quality Objective given in the relevant Daughter Directive, in order to assess the performance of the analyser with respect to the Directives' requirements.

Therefore, a Candidate Method will be treated as a 'black box' measurement system and will undergo testing to determine the uncertainty introduced by all the different performance characteristics of the complete measurement system. These separate uncertainties shall then be combined to give an expanded uncertainty, expressed with a level of confidence of 95%, for the CM, which shall then comply with the measurement uncertainty laid down within the relevant EU Directive.

Testing shall be carried out in two parts. A laboratory test in which two instruments of the same pattern will be tested, and a field test in which these two instruments will be tested together against the relevant Reference Method.

If a CM is a modification to an existing type-approved analyzer, then only the laboratory performance characteristics that are affected by the modification shall be tested and their standard uncertainties calculated. The standard uncertainties associated with unaffected

performance characteristics shall then be used together with these existing standard uncertainties, to determine the combined measurement uncertainty, u_c. An example of such a modification would be a new material for a converter of NO_x to NO within a chemiluminescence NO_x analyzer. Under this modification only the converter efficiency test and the response time test shall be performed in the lab tests. In all cases the field tests shall be performed.

If a CM utilises a measurement method that is different to the EN Standard Method, then all of the tests shall be performed.

The following performance characteristics of the CM will be tested, where applicable:

(i) Laboratory tests

- response time, consisting of rise lag time, rise time, fall lag time and fall time (where applicable);
- laboratory repeatability standard deviation;
- short-term zero and span drift;
- difference between sample port and calibration port (where applicable);
- detection limit ;
- averaging of short-term fluctuations in measurand concentration (where applicable);
- lack of fit (linearity);
- cross-sensitivity to potentially-interfering substances;
- NO_x converter efficiency test (where applicable);
- carry-over (where applicable);
- influence of atmospheric sample pressure and temperature;
- influence of surrounding air temperature
- influence of supply voltage variations.

Both analysers used in the laboratory tests are required to pass all the tests.

(ii) Field tests

- field performance of two CM analysers of the same type(pattern) against the relevant Reference Method to determine whether systematic differences occur in the measured results;
- field repeatability of two CM analysers;
- long-term zero and span drift;
- availability (maintenance interval).

Both analysers used in the field tests are required to pas all the tests.

The performance characteristics calculated from the tests shall be compared to the same performance characteristics defined in Table 8.1.

From the performance characteristics the following standard uncertainties, where applicable, shall be calculated and used to calculate the combined expanded measurement uncertainty of the CM:

Table 8.1. Standard uncertainty components to be included in the combined standard measurement uncertainty

Uncertainty Source	*Symbol*
Repeatability at zero	u_z
Repeatability at 70-80% of the certification concentration	u_s
Between-instrument uncertainty	u_f
Carry over	u_c
Lack of fit (linearity)	u_l

Difference between sample and calibration port	u_a
Effect of short term fluctuations in concentration	u_{av}
Cross sensitivity to interfering substances	u_{H2O}, u_{int}
Variation in sample pressure	u_p
Variation in sample temperature	u_t
Variation in surrounding air temperature	u_{st}
Variation in supply voltage	u_v
NO_x converter efficiency	u_{ce}
Comparison with the Standard Method	U_{CM}
Long-term zero drift	u_{zd}
Long-term span drift	u_{sd}

8.3 *Definitions applicable to automatic measurement systems*

8.3.1 Independent measurement — An individual measurement that is not influenced by a previous individual measurement, by separating two individual measurements by at least four response times.

8.3.2 Individual measurement — A measurement averaged over a time period equal to the response time of the analyser.

8.4 *Laboratory tests*

8.4.1 Test concentrations

Laboratory tests are performed, in principle, over the range of concentrations specified in the EN Standard for the reference method. A more restricted certification range may be selected by a Member State if judged appropriate. (The CM will then only be applicable to results obtained in this restricted certification range.) Test concentrations specified here for the laboratory tests are based on the maximum of the selected certification range, unless specified otherwise.

8.4.2 Response time

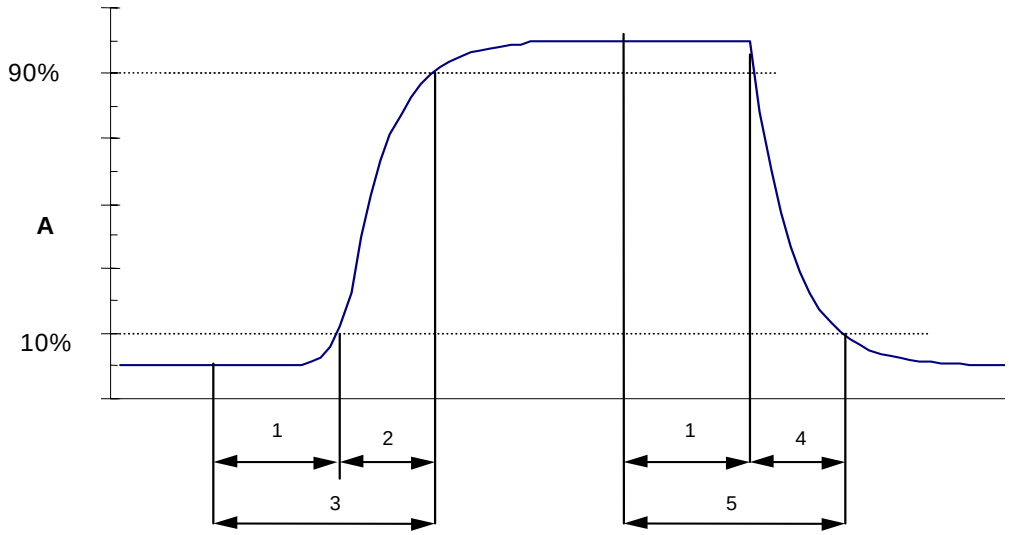
The tests of response time shall be performed on all Candidate Methods that give approximately real-time measurements.

For instruments such as gas chromatographs, the concept of response time is inappropriate and the response time shall be taken to be the time interval of the sampling. For example, if the instrument collects a sample once an hour, then the assumed response time will be one hour. The requirement on response time for this type of CM is that the assumed response time is less than 25% of the required averaging period.

For real-time CMs the following test procedure shall be used:

Apply a step change of gas concentration to the CM. The step change value shall be within 20% and 80% of the maximum of the certification range of the CM and will normally be to 80%, and need not start from or end at zero concentration. The applied change of concentration must reach 90% of the full change of value within 10 seconds of the start of the change. Record the response of the CM to the applied step change of concentration and determine the lag time (the time taken by the CM to indicate 10% of the final concentration value), the response time (the time taken for the CM to indicate 90% of the final concentration value) and the rise time (the difference between the response time and the lag time). Measurements will be made with at least four step changes of concentration in both positive and negative directions. As well as calculating the individual rise and fall response times the relative difference in response times will also be calculated.

When the reading of 98 % of the applied concentration has been reached, the span gas can be changed to zero again. This event is the start (t = 0) of the (fall) lag time. When the reading of 2 % of the applied concentration has been reached the whole cycle as shown in Figure 8.1 is complete.



Key

A Analyser response
1 Lag time
2 Rise time
3 Response time (rise)
4 Fall time
5 Response time (fall)

Figure 8.1 — Diagram illustrating the response time

The requirement on response time being less than 25% of the required averaging period shall be used. An additional requirement is imposed of up to 10% in the relative difference between response rise time and response fall time.

For CMs measuring NO and NO_2 simultaneously, the response time shall be determined for both NO and NO_2 test gases.

Where the CM uses an adaptive filter for data smoothing, the response times of the CM shall be measured with both the filter enabled and disabled. The maximum response time measured in both these sets of tests shall be compared with the performance standard.

8.4.3 Short–term drift

The CM is calibrated at both zero and at 70% to 80% of the maximum of the certification range and adjusted as appropriate. It is then supplied with test gas at zero concentration, after the period equivalent to one independent reading has passed, 20 individual measurements will be recorded of the CM's output. The CM is then supplied with test gas at a concentration around

70% to 80% of the maximum of the certification range and the equivalent measurements recorded.

The CM shall be operated under the laboratory conditions whilst analysing ambient air. After a minimum period of 12 hours the repeatability test is repeated. The averaged values obtained for zero and 70% to 80% of the maximum of the certification range shall be calculated. This test shall be used to show that the 12-hour drift is not the dominant factor in any of the test results.

The short-term drift at zero and at 70% to 80% of the maximum of the certification range shall be calculated as follows:

$$D_{s,z} = \left(C_{z,2} - C_{z,1}\right) \quad (8.1)$$

where

$D_{s,z}$ = the 12-hour drift at zero;

$C_{z,1}$ = the average of the zero gas measurements at the beginning of the drift period (just after calibration);
$C_{z,2}$ = the average of the zero gas measurements at the end of the drift period (12 hours).

D_{sz} shall comply with the performance criterion for short term drift at zero given in the relevant EN Standard for the measurand.

$$D_{s,s} = \left(C_{s,2} - C_{s,1}\right)$$

Span drift now zero drift corrected

$$D_{s,s} = \left(C_{s,2} - C_{s,1}\right) - D_{s,z} \quad (8.2)$$

where:
$D_{s,s}$ = the 12-hour drift at the test concentration C_t (nmol/mol);
$C_{s,1}$ = the average of the span gas measurements at the beginning of the drift period (just after calibration) (nmol/mol);
$C_{s,2}$ = the average of the span gas measurements at the end of the drift period (12 hours) (nmol/mol)

$D_{s,s}$ shall comply with the performance criterion for short term drift at span levels given in the relevant EN Standard for the compound under investigation.

8.4.4 Repeatability for continuous measuring CMs

Test gases shall be supplied to the CM at zero concentration and the highest numerical limit or target value specified for the pollutant for a period equivalent to one independent measurement, and then 20 individual measurements of the CM's output are recorded.
From these measurements the repeatability standard deviation (s_i) at zero concentration and at concentration c_t shall be calculated according to:

$$s_i = \sqrt{\frac{\sum \left(y_i - \bar{y}\right)^2}{n - 1}} \quad (8.3)$$

where:

s_i = the repeatability standard deviation;
y_i = the ith measurement;
$\bar{y}$ = the average of the 20 measurements;
n = the number of measurements, n =20.

The repeatability standard deviation shall be calculated separately for both series of measurements (zero gas and concentration c_t) and the repeatability (r_i) is calculated according to:

$$r_i = t_{n\ 1,0.05} \cdot s_i \qquad (8.4)$$

where:
$t_{n\text{-}1,\ 0,05}$ = the two-sided Students t-factor at a confidence level of 0,05, with n-1 degrees of freedom (for n = 20, $t_{n\text{-}1,\ 0,05}$ = 2,09);
s_l = the repeatability standard deviation.

r_l shall comply with the performance criteria for repeatability at zero and repeatability at the limit value concentration, respectively, given in the EN Standard for the measurand.

The standard uncertainties u_z and u_s, for repeatability are equal to the repeatability standard deviation, $s_{l,}$ calculated above, for the zero and the limit/target value concentrations.

8.4.5 Carry over and repeatability for CMs collecting samples onto a sorbent prior to analysis

CMs that collect samples by absorption or other similar means and then subsequently analyse them shall be tested for the carry-over of measurand from one sample to the next. The CM shall be supplied with test gas for one sampling period, at the highest numerical limit or target value specified for the pollutant, followed by one sampling period of zero gas. This procedure shall be repeated 20 times and the results shall be used to calculate both the repeatability standard deviation at the limit value concentration and the carry over standard deviation according to:

$$s_i = \sqrt{\frac{\sum \left(y_i - \bar{y}\right)^2}{n - 1}} \qquad (8.5)$$

where:
s_i = the repeatability / carry over standard deviation;
y_i = the ith measurement ;
$\bar{y}$ = the average of the 20 measurements;
n = the number of measurements, n =20.

The repeatability standard deviation shall be calculated separately for both series of measurements (zero gas and concentration c_t) and the repeatability (r_i) is calculated according to eq. (8.4).

r_l shall comply with the performance criteria for carry over and repeatability specified in the relevant EN Standard, respectively.

The standard uncertainties u_c and u_s, for carry over and repeatability are equal to the repeatability standard deviation, s_i calculated above, for the carry over and limit value test concentrations.

8.4.6 Lack of fit (linearity)

The CM shall be adjusted at a concentration of about 90 % of the maximum of the certification range. The linearity of the CM shall then be tested over the range 0 % to 95 % of the maximum of

the certification range of the CM using at least 6 concentrations (including the zero point). At each concentration (including zero) at least 5 independent readings shall be performed.

The concentrations shall be applied in the following sequence: 80 %, 40 %, 0 %, 60 %, 20 % and 95 %. The dilution ratios for the applied concentrations shall be less than 1,5 % with respect to each other.

Continuous measuring CMs

After each change in concentration a delay of at least 4 response times shall be taken into account before the next measurement is performed.

Non-continuous measuring CMs

After each change in concentration at least a delay of 1 response time shall be taken into account before the next measurement is performed.

For both, a linear regression function is calculated from the measured mean responses at each concentration. The relative residual for each measured concentration is calculated. The largest relative residual ρ_{max} and the actual residual at the zero concentration will be compared against the performance criteria for lack of fit.

The standard uncertainty due to the lack of fit at the limit value concentration, u_l, is calculated according to:

$$u_l = \frac{\rho_{max} \cdot LV}{\sqrt{3}} \qquad (8.6)$$

where:
u_l = the standard uncertainty due to lack of fit at the limit value concentration;
LV = the highest numerical limit or target value specified for the pollutant.

8.4.7 Differences between sample and calibration port

If the CM has different ports for sample gas and calibration gas, the difference in response of the CM when test gas is introduced through the sample or calibration port shall be tested. The test shall be carried out by supplying test gas, with a concentration of 70 % to 80 % of the maximum of the certification range, through the sample port for a time period equal to one independent measurement. Three individual measurements shall then be taken of the CM output. Zero gas is then supplied to the sample port for a time period equal to one independent measurement. This test is repeated with the teat gas supplied to the calibration port instead of the sample port. A delay of 4 response times should be left between testing the sample and calibration ports to allow for flushing.

The difference shall be calculated according to:

$$D_{sc} = \left| \frac{y_s - y_c}{c_t} \right| \times 100\% \qquad (8.7)$$

where

D_{sc} = the difference sample/calibration port (%);

y_s = the average of the concentrations measured using the sample port;

y_c = the average of the concentrations measured using the calibration port;

c_t = the concentration of the test gas.

D_{sc} shall comply with the performance criterion for the difference between the sample and calibration port in the relevant EN Standard for the compound under investigation.

The standard uncertainty due to the difference between the sample and calibration port, u_a, is calculated according to:

$$u_a = \frac{\frac{D_{sc}}{100} \cdot LV}{\sqrt{3}} \qquad (8.8)$$

where
u_a = the standard uncertainty due to the difference between the sample and calibration ports, at the highest numerical limit or target value specified for the pollutant.

8.4.8 Effect of short-term fluctuations in concentration (averaging test)

The averaging test gives a measure of the uncertainty in the averaged values caused by concentration variations in the sampled air shorter than the time scale of the measurement process in the analyser. For the determination of the uncertainty due to the averaging the following concentrations are applied to the analyser and readings are taken at each concentration:

1) CO, O_3, SO_2 and benzene measuring CMs

- a constant concentration of the measurand at a concentration that is about twice the highest numerical limit value specified for the pollutant
- a stepwise varied concentration of the measurand between zero and about twice the highest numerical limit value specified for the pollutant.

2) CMs measuring NO and NO_2 simultaneously

- a constant concentration of NO_2 at a concentration, which is about the hourly limit value;
- a stepwise varied concentration of NO between zero and a concentration corresponding to about six times the hourly limit value for NO_2.

For non-continuously measuring CMs the time period (t_c) of the constant concentration shall be at least equal to a period necessary to obtain two cycle periods (which equals to at least two response times). The time period (t_v) of the varying concentration shall be at least equal to to a period necessary to obtain four cycle periods (which equals to at least four response times). The time period (t_D) for the measurand concentration shall be 90 seconds followed by a period (t_{zero}) of 90 seconds of zero concentration.

The change from t_D to t_{zero} shall be within 0,5 seconds. The change from t_C to t_V shall be within one response time of the analyser under test.
The averaging effect (X_{av}) is calculated according to:

$$X_{av} = \frac{C_{const}^{av} - 2 \times C_{var}^{av}}{C_{const}^{av}} \times 100\% \qquad (8.9)$$

where
X_{av} = the averaging effect (%);

$C_{\text{const}}^{\text{av}}$ = the average of the at least 4 independent measurements during the constant concentration period (t_c);

$C_{\text{var}}^{\text{av}}$ = the average of the at least 4 independent measurements during the variable concentration period (t_v).

X_{av} shall comply with the requirements for the measurand in the relevant EN Standard Method. For instruments measuring NO and NO2 simultaneously, Xav shall be calculated for both channels and compared with these requirements.

The resulting uncertainty u_{av} is calculated as

$$u_{av} = \frac{X_{av}}{100\sqrt{3}} \tag{8.10}$$

8.4.9 Variation in sample-gas pressure

Measurements are taken at a concentration of about 70 % to 80 % of the maximum of the certification range of the CM at an absolute pressure of about 80 kPa ± 0,2 kPa and at an absolute pressure of about 110 kPa ± 0,2 kPa. At each pressure the test gas is supplied for a time period equal to one independent measurement and then three individual measurements will then be taken of the CMs output. From these measurements the averages at each pressure are calculated.

Measurements at different pressures shall be separated by at least 4 response times for continuous measuring CMs and one response time for non-continuous measuring CMs.

The sensitivity coefficient for the influence of sample gas pressure is calculated by:

$$\frac{\Delta C}{\Delta P} = \left| \frac{C_{P2} - C_{P1}}{P_2 - P_1} \right| \tag{8.11}$$

where

C_{P_1} = the average concentration of the measurements at sampling gas pressure P_1;

C_{P_2} = the average concentration of the measurements at sampling gas pressure P_2;

P_1 = the sampling gas pressure P_1 (kPa);

P_2 = the sampling gas pressure P_2 (kPa).

The test parameter b_{gp} to be compared to the test criterion in the relevant EN Standard for the compound investigated is then calculated as

$$b_{gp} = \frac{\Delta C}{\Delta P} \cdot \frac{100}{C_t} \tag{8.12}$$

where C_t is the applied test gas concentration.

The standard uncertainty due sample pressure variation, u_p, is calculated according to:

$$u_p = \frac{\Delta C}{\Delta P} \cdot \left| \frac{P_{max} - P_{min}}{\sqrt{3}} \right| \tag{8.13}$$

where P_{max} and P_{min} are the extremes of pressures encountered in practice.

For calculation of the standard uncertainty from the results of the type-approval test, P_{max} and P_{min} equal P_2 and P_1, respectively.

8.4.10 Variation in sample-gas temperature

For the determination of the dependence of the sample gas temperature measurements shall be performed at sample gas temperatures of 273K and 303K. The temperature dependence shall be determined at a concentration of about 70 % to 80 % of the maximum of the certification range of the CM. At each temperature the test gas is supplied for a time period equal to one independent measurement and then three individual measurements will then be taken of the CMs output.

The sample gas temperature, measured at the inlet of the analyser, shall be held constant for at least 30 minutes before any measurements are taken.

The sensitivity coefficient for the influence of sample gas temperature is calculated as:

$$\frac{\Delta C}{\Delta T_{sg}} = \left| \frac{C_{T2} - C_{T1}}{T_2 - T_1} \right| \tag{8.14}$$

where

C_{T_1} = the average concentration of the measurements at sample gas temperature T_1 (°C);

C_{T_2} = the average concentration of the measurements at sample gas temperature T_2 (°C);

T_1 = the sample gas temperature T_1 (°C);

T_2 = the sample gas temperature T_2 (°C).

The test parameter b_{gp} to be compared to the test criterion in the relevant EN Standard for the compound investigated is then calculated as

$$b_{sg} = \frac{\Delta C}{\Delta T_{sg}} \cdot \frac{100}{C_t} \tag{8.15}$$

where C_t is the applied test gas concentration.

The standard uncertainty due sample pressure variation, u_p, is calculated according to:

$$u_p = \frac{\Delta C}{\Delta T_{sg}} \cdot \left| \frac{T_{sg,max} - T_{sg,min}}{\sqrt{3}} \right| \tag{8.16}$$

where $T_{sg,max}$ and $T_{sg,min}$ are the extremes of temperature encountered in practice.

For calculation of the standard uncertainty from the results of the type-approval test, P_{max} and P_{min} equal T_2 and T_1, respectively.

8.4.11 Surrounding temperature variation

The influence of the surrounding air temperature shall be determined at the following temperatures (within the specifications of the manufacturer):

- the minimum specified temperature ($T_{e,min}$);
- at the laboratory temperature (T_1);
- at the maximum specified temperature ($T_{e,max}$).

For these tests a climate chamber is necessary.

The influence shall be determined at a concentration around 70 % to 80 % of the maximum of the certification range of the CM. At each temperature the test gas is supplied for a time period equal to one independent measurement and then three individual measurements will then be taken of the CM's output. At each temperature measurements at zero and a concentration around 70 % to 80 % of the maximum of the certification range of the CM shall be performed.

At each temperature setting the criteria for warm-up or stabilisation time are to be met.
The measurements shall be performed in the following sequence of the temperature settings:
T_1, $T_{e,min}$, T_1 and T_1, $T_{e,max}$, T_1

At the first temperature (T_1) the CM shall be adjusted at zero and at span level (70 % to 80 % of the maximum of the certification range). Then measurements shall be carried out at T_1, at $T_{e,min}$ and again at T_1. This procedure shall be repeated at the temperature sequence of T_1 ,$T_{e,max}$, and at T_1.

In order to exclude any possible drift due to factors other than temperature, the measurements at T_1 are averaged, which is taken into account in the following formula for calculation of the sensitivity coefficient for the influence of surrounding temperature:

$$\frac{\Delta C}{\Delta T_e} = \left| \frac{y_T - \frac{y_1 + y_2}{2}}{T - T_I} \right|^{\frac{1}{2}} \qquad (8.17)$$

where
y_T = the average of the measurements at $T_{e,min}$ or $T_{e,max}$;
y_1 = the first average of the measurements at T_1 just after calibration;
y_2 = the second average of the measurements at T_1 just before calibration;
T_I = the surrounding air temperature at the laboratory (°C);
T = the surrounding air temperature $T_{e,min}$ or $T_{e,max}$ (°C).

For reporting the surrounding air temperature dependence the higher value is taken of the two calculations of the temperature dependence at $T_{e,min}$ and $T_{e,max}$.

This value is then compared to the test criterion in the relevant EN Standard for the compound investigated.

The standard uncertainty due to surrounding temperature variation, u_{st}, is calculated according to:

$$u_{st} = \frac{\Delta C}{\Delta T_e} \cdot \left| \frac{T_{e,max} - T_{e,min}}{\sqrt{3}} \right| \qquad (8.18)$$

where $T_{e,max}$ and $T_{s,min}$ are the extremes of surrounding temperature encountered in practice.

For calculation of the standard uncertainty from the results of the type-approval test, $T_{e,eax}$ and $T_{e,min}$ are the temperatures used in this test.

8.4.12 Variation due to supply voltage

The influence of changes in the electrical supply voltage shall be determined at both ends of the specified voltage range at zero concentration and at a concentration around 70 % to 80 % of the maximum of the certification range of the CM. At each voltage the test gas is supplied for a time

period equal to one independent measurement and then three individual measurements will then be taken of the CMs output. At each voltage measurements at zero and around 70 % to 80 % of the maximum of the certification range shall be performed.

The sensitivity coefficient for the influence of voltage dependence is calculated according to:

$$\frac{\Delta C}{\Delta V} = \left| \frac{C_{V2} - C_{V1}}{V_2 - V_1} \right| \tag{8.19}$$

where

C_{V_1} = the average concentration reading of the measurements at voltage V_1;

C_{V_2} = the average concentration reading of the measurements at voltage V_2;

V_1 = the minimum voltage V_{min} (V) specified by the manufacturer;
V_2 = the maximum voltage V_{max} (V) specified by the manufacturer.

For reporting the dependence on voltage the highest value of the result at zero and span level shall be taken. This value is then compared to the test criterion in the relevant EN Standard for the compound under investigation.

For an analyser operating on direct current the type approval test of voltage variation shall be carried out over the range of ± 10 % of the nominal voltage.

The standard uncertainty due to voltage variation, u_v, shall be calculated according to:

$$u_v = \frac{\Delta C}{\Delta V} \cdot \left| \frac{V_{max} - V_{min}}{\sqrt{3}} \right| \tag{8.20}$$

where V_{max} and V_{min} are the extremes of line voltage encountered in practice.

For calculation of the standard uncertainty from the results of the type-approval test, V_{max} and V_{min} equal V_2 and V_1, respectively.

8.4.13 Cross-sensitivity to interfering substances

The analyser's response to certain interfering substances, which are to be expected to be present in ambient air and which may also interfere with the CMs measurement process shall be tested. These interferents can give a positive or negative response. The test shall be performed at zero and at a test concentration (c_t) similar to the highest numerical limit or target value specified for the pollutant.
The concentration of the mixtures of the test gases with the interferent shall have an uncertainty of less than 5 % and shall be traceable to National Standards. The influence of each interferent shall be determined separately. A correction on the concentration of the measurand shall be made for the dilution effect due to addition of an interferent (e.g. water vapour).

After adjustment of the analyser at zero and 70% to 80% of the maximum of the certification range the analyser shall be fed with a mixture of zero gas and the interferent to be investigated. This mixture will be supplied for a time period equal to one independent measurement, and, following this, three individual measurements will then be taken of the CMs output. This procedure shall be repeated with a mixture of the measurand at concentration c_t and the interferent to be investigated. The influence quantity at zero and concentration c_t are calculated from:

$$Y_{int,z} = y_z \quad (8.21)$$

where
$Y_{int,z}$ = the influence quantity of the interferent at zero;
y_z = the average of the measurements at zero.
and:

$$Y_{int,c_t} = y_{c_t} - c_t \quad (8.22)$$

where
Y_{int,c_t} = the influence quantity of the interferent at concentration c_t;
y_{c_t} = the average of the measurements at concentration c_t;
c_t = the concentration of the applied gas at the level of the hourly limit value.

The standard uncertainties due to interfering substances, u_{H2O} and u_{int}, are calculated as follows.

The influence quantity of water vapour is established at a water concentration of 16 mmol/mol. The uncertainty, however, is to be established at a water concentration of 18 mmol/mol. The standard uncertainty due to interference by the presence of water vapour at the highest numerical limit value specified for the pollutant, u_{H2O}, is therefore calculated according to:

$$Y_{H_2O,z,max} = (18/16)Y_{H_2O,z} \quad (8.23)$$

$$Y_{H_2O,c_t,max} = (18/16)Y_{H_2O,c_t} \quad (8.24)$$

$$Y_{H_2O,max} = ((Y_{H_2O,c_t,max} - Y_{H_2O,z,max})/c_t) \cdot LV + Y_{H_2O,z,max} \quad (8.25)$$

$$u_{H_2O} = \left|Y_{H_2O}/c_{H_2O,max}\right| \cdot \sqrt{(c_{H_2O,max}{}^2 + c_{H_2O,max}c_{H_2O,min} + c_{H_2O,min}{}^2)/3} \quad (8.26)$$

where
$Y_{H_2O,z,max}$ = the influence quantity of an H_2O concentration of 18 mmol/mol at zero concentration of the measurand (nmol/mol);

$Y_{H_2O,z}$ = the influence quantity of an H_2O concentration of 16 mmol/mol at zero concentration of the measurand (nmol/mol);

$Y_{H_2O,c_t,max}$ = the influence quantity of an H_2O concentration of 18 mmol/mol at the test concentration c_t of the measurand (nmol/mol);

Y_{H_2O,c_t} = the influence quantity of an H_2O concentration of 16 mmol/mol at the test concentration c_t of the measurand (nmol/mol);

Y_{H_2O} = the influence quantity of an H_2O concentration of 18 mmol/mol at the hourly limit value (nmol/mol);

c_t = the test gas concentration of the measurand;

LV = the highest numerical limit value specified for the pollutant;

u_{H_2O} = the standard uncertainty due to interference by the presence of water vapour (nmol/mol);

$c_{H_2O,max}$ = the maximum concentration of water vapour (mmol/mol) (= 18 mmol/mol);

$c_{H_2O,min}$ = the minimum concentration of water vapour (mmol/mol) (= 6 mmol/mol).

The standard uncertainty due to each interfering compound (other than water vapour) at the highest numerical limit value specified for the pollutant, u_{int}, is calculated according to:

$$Y_{int} = ((Y_{int,c_t} - Y_{int,z}) / c_t) \cdot LV + Y_{int,z} \quad (8.27)$$

$$u_{int} = \left| Y_{int} / c_{int,max} \right| \cdot \sqrt{(c_{int,max}^{\ 2} + c_{int,max} c_{int,min} + c_{int,min}^{\ 2}) / 3} \quad (8.28)$$

where

Y_{int,c_t} = the influence quantity of the maximum concentration of the relevant interfering compound at the test concentration c_t of the measurand;

$Y_{int,z}$ = the influence quantity of the maximum concentration of the relevant interfering compound at zero concentration of the measurand;

Y_{int} = the influence quantity of the relevant interfering compound;

c_t = the test concentration of the measurand at the level of the hourly limit value;

LV = the highest numerical limit value specified for the pollutant;

u_{int} = the standard uncertainty due to interference by the presence of a chemical compound;

$c_{int,max}$ = the maximum concentration of interfering compound;

$c_{int,min}$ = the minimum concentration of interfering compound.

According to ISO 14956 the summed uncertainties due to the interferents with positive impact and the summed uncertainties of the interferents with negative impact shall be calculated according to:

$$S_{u_{int,pos}} = u_{int,1,pos} + u_{int,2,pos} + ... + u_{int,n,pos} \quad (8.29)$$

$$S_{u_{int,neg}} = u_{int,1,neg} + u_{int,2,neg} + ... + u_{int,n,neg} \quad (8.30)$$

Take the highest sum as the representative value for all interferents.

$$u_{int,pos} = \sqrt{(u_{int,1,pos} + u_{int,2,pos} + ... + u_{int,n,pos})^2} \quad (8.31)$$

$$u_{int,neg} = \sqrt{(u_{int,1,neg} + u_{int,2,neg} + ... + u_{int,n,neg})^2} \quad (8.32)$$

where

$u_{int,pos}$ = the sum of uncertainties due to interferents with positive impact;

$u_{int,1,pos}$ = the uncertainty due to the 1st interferent with positive impact ;

$u_{int,n,pos}$ = the uncertainty due to the *n*th interferent with positive impact;

$u_{int,neg}$ = the sum of uncertainties due to interferents with negative impact ;

$u_{int,1,neg}$ = the uncertainty due to the 1st interferent with negative impact;

$u_{int,n,neg}$ = the uncertainty due to the *n*th interferent with negative impact.

8.4.14 NO_2 converter efficiency

The converter efficiency is determined by measurements with calculated amounts of NO_2. This can be achieved by means of gas-phase titration of NO to NO_2 with ozone.

The test is to be performed at two concentration levels: at 50 % and 95 % of the maximum of the certification range of NO_2.

The NO_x analyser shall be calibrated on the NO and NO_x channel with a NO concentration around 70 % to 80 % of the maximum of the certification range of NO. Both channels shall be set to read the same value and the values shall be recorded.

A known concentration of about 50 % of the maximum of the certification range of NO shall be supplied to the analyser until a stable output signal is achieved. This stabilisation period shall be at least four times the response time of the analyser. Four individual measurements are taken at the NO and NO_x channel. The NO will then be reacted with O_3 to produce the required concentration of NO_2, the NO residue after the gas phase titration reaction shall be 10 % to 20 % of the original NO concentration. This mixture with a constant NO_x concentration shall be supplied to the analyser until a stable output signal is achieved. This stabilisation period shall be at least four times the response time of the analyser.

Four individual measurements are then taken at the NO and the NO_x channel. The O_3 supply shall be switched off and the analyser supplied with only NO until a stable output signal is achieved. This stabilisation period shall be at least four times the response time of the analyser. Then the average of the four individual measurements at the NO and NO_x channel is checked to see whether it is equal within 1 % of the original values.

Repeat the test with a NO test concentration of approximately 95% of the maximum of the certification range of NO_2.

Calculate the converter efficiency from:

$$E_{conv} = \left(1 - \frac{(NO_X)_i - (NO_X)_f}{(NO)_i - (NO)_f}\right) \times 100\,\% \qquad (8.33)$$

where

E_{conv} = the converter efficiency in %;

$(NO_x)_i$ = the average of the four individual measurements at the NO_x channel at the initial NO_x concentration;

$(NO_x)_f$ = the average of the four individual measurements at the NO_x channel at the resulting NO_x concentration after applying O_3;

$(NO)_i$ = the average of the four individual measurements at the NO channel at the initial NO concentration;

$(NO)_f$ = the average of the four individual measurements at the NO channel at the resulting NO concentration after applying O_3.

The lowest value of the two converter efficiencies shall be used to calculate the standard uncertainty due to converter efficiency, u_{ce}, according to:

$$u_{ce} = \frac{E_{conv}}{100\sqrt{3}} \cdot LV \qquad (8.34)$$

where LV is the highest numerical limit value specified for the pollutant.

8.5 *Field test*

8.5.1 General

In the field test two CMs of the same type (pattern) are tested for availability (period of unattended operation), between-instrument uncertainty and long-term drift. The CMs are run in parallel at one and the same sampling point at a selected monitoring station. Operational requirements are given below for the correct determination of the long-term drift and the between-instrument uncertainty.

The EN Standard Method will be operated alongside the two CMs, with parallel measurements from one and the same sampling point. From these results any systematic differences between the results obtained by the CM and the Standard Method will be determined.

8.5.2 Experimental conditions

Test sites shall be representative for typical conditions for which equivalence will be claimed, including possible episodes of high concentrations. A minimum of 4 comparisons shall be performed with particular emphasis on the following variables, if appropriate:

- Composition of the air, notably high and low concentrations of the measured compound and potential interferents
- Air humidity and temperature (high and low) to cover any effects on sampling efficiency or desorption efficiency
- Wind speed (high and low) to cover any dependency of sampler performance due to deviations from ideal behaviour.

Each comparison shall cover a minimum of one month of uninterrupted monitoring during which hourly-average measurement results shall be collected.

Both methods shall be operated under conditions reflecting practical application in the field, e.g., calibration intervals, appropriate span and zero checks. At the beginning of the test both CMs will be adjusted to read the same value.

During the tests, the following information shall be collected and recorded

- calibration procedures, equipment and intervals
- (results of) quality checks
- temperature and pressure of the sampled air
- other conditions relevant for the measurements performed (e.g., air humidity)
- particular events/situations that may be of influence on measurement results.

8.5.3 Evaluation of data collected

8.5.3.1 *Suitability of datasets*

Of the full dataset at least 20% of the results shall be greater than or equal to the upper assessment threshold specified in the relevant Daughter Directive.
Data shall only be removed from the data set when sound technical reasons can be found for doing so. All valid data shall be used for further evaluation.

Note: Indications of outlying data (pairs) may be obtained using Grubb's tests on the individual single-period variances. Outlier tests are to be performed at the 99% level.

8.5.3.2 *Between-instrument uncertainty*

The relative between-sampler/instrument uncertainty w_{bs} is calculated from the differences of all hourly results of the candidate samplers/instruments operated in parallel as:

$$w_{bs}^2 = \frac{\sum_{i=1}^{n} (y_{i,1} - y_{i,2})^2}{2n\bar{y}^2} \qquad (8.35)$$

where
$y_{i,1}$ and $y_{i,2}$ are the results of parallel measurements for a single 1-hour period i
n = number of hourly measurement results
$\bar{y}$ = average of all measurement results of the Candidate Method.

The between-instrument uncertainty shall comply with the criterion given in Annex A.

8.5.3.3 *Comparison with the Standard Method*

For a comparison with the Standard Method the results of the parallel measurements of reference samplers/instruments and candidate samplers/instruments are averaged to give one result x_i or y_i for common measurement periods of equal length.

For the evaluation of the uncertainty due to the 'lack of comparability' between Candidate and Standard Methods it is assumed that the relationship between measurement results of both methods can be described by a linear relation of the form

$$y_i = a + bx_i \qquad (8.36)$$

The relation between the average results of the Candidate Method and those of the Standard Method is established using a regression technique that leads to a symmetrical treatment of both variables. A commonly applied technique is orthogonal regression [7,28].

The uncertainty due to lack of comparability will be a function of the concentration of the measurand. The general relationship describing the dependence of $u_{C\text{-}S}$ on x_i is given by

$$u_{C\_S}^2(y_i) = \frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)x_i]^2 \qquad (8.37)$$

where
RSS = the sum of (relative) residuals resulting from the orthogonal regression
$u(x_i)$ = random uncertainty of the Standard Method; as such, the value of u_{bs} calculated for the application of the Standard Method in these tests may be used. (See also note of equation 7.30)

Algorithms for the calculation of *a* and *b* and their variances are given in Annex B.

RSS, the sum of (relative) residuals is calculated using eq. 8.38a or 8.38b, depending on whether the reiduals or relative residuals are constant.

$$RSS = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \text{ when } (y_i - a - bx_i)^2 \text{ is constant} \qquad (8.38a)$$

$$RSS = (a + bx_i)^2 \sum_{i=1}^{n} \left(\frac{y_i}{a + bx_i} - 1 \right)^2 \text{ when } \left(\frac{y_i}{a + bx_i} - 1 \right)^2 \text{ is constant} \quad (8.38b)$$

8.5.3.3 *Calculation of the combined 'field' uncertainty to be assigned to the Candidate Method*

The combined relative field uncertainty of the Candidate Method $w_{CM,field}$ is calculated by combining the contributions found in 8.5.3.1 and 8.5.3.2 as follows:

$$w^2_{CM,field}(y_i) = \frac{u^2_{C\_S}(y_i)}{y_i^2} \quad (8.39)$$

The uncertainty at the limit value with the shortest averaging period, $w_{CM,field}$, is calculated by taking as y_i the concentration at the limit value.

8.5.3.4 *Calculation of the expanded field uncertainty of Candidate Method*

For each of the datasets the expanded relative uncertainty of the results of the Candidate Method is calculated by multiplying $w_{c,CM}$ by a coverage factor k reflecting the appropriate number of degrees of freedom resulting from the determination of $w_{c,CM}$ as

$$W_{CM,field} = k \cdot w_{CM,field} \quad (8.40)$$

In view of the large number of experimental results available, a coverage factor k=2 can be used.

8.5.3.5 *Long-term drift*

After each bi-weekly calibration the drift of the analysers under test must be calculated at zero and at span following the procedures as given below. If the drift compared to the initial calibration exceeds one of the performance criteria for drift at zero or span level, the "period of unattended operation" equals the number of weeks till the observation of the infringement, minus two weeks. For uncertainty calculations the values for "long term drift" are the values for zero and span drift over the period of unattended operation.

The long-term drift is calculated as follows:

$D_{L,z} = (C_{z,2} - C_{z,1})$ (8.41)

where
$D_{L,z}$ = the drift at zero;
$C_{z,1}$ = the average of five individual zero gas measurements at the beginning of the drift period (just after the initial calibration);
$C_{z,2}$ = the average of five individual zero gas measurements at the end of the drift period (without any mathematical correction applied to the data).

$$D_{L,s} = \frac{(C_{s,2} - C_{s,1}) - D_{L,z}}{C_{s,1}} \times 100\% \quad (8.42)$$

where
$D_{L,,s}$ = the drift at span concentration, c_t;

$C_{s,1}$ = the average of five individual span gas measurements at the beginning of the drift period (just after the initial calibration);
$C_{s,2}$ = the average of five individual span gas measurements at the end of the drift period (without any mathematical correction applied to the data).

The standard uncertainty due to long-term zero drift, u_{zd}, is calculated according to:

$$u_{dz} = \frac{D_{L,z}}{\sqrt{12}} \tag{8.43}$$

The standard uncertainty due to long-term span drift, u_{sd}, is calculated according to:

$$u_{sd} = \frac{\frac{D_{L,s}}{100} LV}{\sqrt{12}} \tag{8.44}$$

where *LV* is the hourly limit value.

8.5.3.6 *Period of Unattended Operation*

The period of unattended operation is the time period within which the drift is within the performance criterion for long term drift. If the manufacturer specifies a shorter period for maintenance, then this will be taken as the period of unattended operation. If one of the analysers malfunctions during the field test, then the field test shall be restarted to show whether the malfunction was coincidental or bad design.

8.5.3.7 *Period of availability of the analyser*

The correct operation of the CM shall be checked at least every 14 days. It is recommended to perform this check every day during the first 14 days. These checks consists of plausibility checks on the measured values, as well as when available status signals and other relevant parameters. Time, duration and nature of any malfunctioning shall be logged.

The total time period with useable measuring data is the period during the field test during which valid measuring data of the ambient air concentrations are obtained. In this time period the time needed for calibrations, conditioning of sample lines, filters and maintenance shall not be included.

The availability of the analyser is calculated as:

$$A_{\mathrm{a}} = \frac{t_{\mathrm{u}}}{t_{\mathrm{t}}} \times 100\,\% \tag{8.45}$$

where
A_{a} = the availability of the CM;
t_{u} = the total time period with validated measuring data;
t_{t} = the time period of the field test minus the time for calibration, conditioning and maintenance.

t_{u} and t_{t} shall be expressed in the same units (e.g. hours).

8.6 *Determination of the combined measurement uncertainty*

The standard uncertainties from Table 1, where applicable, are combined by the sum of squares method to give the combined standard measurement uncertainty, u_c, according to:

$$u_{CM,lab} = \sqrt{u_z^2 + u_r^2 + u_c^2 + u_l^2 + u_a^2 + u_{av}^2 + u_{H2O}^2 + u_{\text{int}}^2 + u_p^2 + u_t^2 + u_{st}^2 + u_v^2 + u_{ce}^2 + u_{bs}^2 + u_{zd}^2 + u_{sd}^2}$$

(8.46)

The following are CM specific:

u_z will only be included for continuous measuring CMs;
u_c will only be included for non-continuous measuring CMs;
u_{cv} will only be included for NO_x measuring CMs that use a converter to convert NO_x to NO.

From $u_{CN,lab}$, the relative uncertainty at the limit value is calculated as

$$w_{CM,lab} = \frac{u_{CM,lab}}{LV} \qquad (8.47)$$

where LV is the hourly limit value of the measurand.

8.7 *Calculation of the expanded laboratory uncertainty of Candidate Method*

The expanded relative 'laboratory' uncertainty of the results of the Candidate Method is calculated by multiplying $w_{CM,lab}$ by a coverage factor k reflecting the appropriate number of degrees of freedom resulting from the determination of $w_{CM,lab}$ as

$$W_{CM,lab} = k \cdot w_{CM,lab} \qquad (8.48)$$

In view of the large number of experimental results available, a coverage factor k=2 can be used.

8.8 *Evaluation of test results*

The resulting uncertainty estimates $W_{CM,lab}$ and $W_{CM,field}$ are intercompared and compared with the expanded relative uncertainty based on the data quality objective for the Standard Method W_{dqo}. Here, 3 situations may occur.

1. $W_{CM,lab} > W_{dqo}$: the Candidate Method is not accepted as an equivalent method

2. $W_{CM,lab} \leq W_{dqo}$ and $W_{CM,field} > W_{CM,lab}$: the Candidate Method is accepted conditionally; before final acceptance, the uncertainty evaluation from the laboratory tests should be re-evaluated and corrected such that situation 3 occurs

3. $W_{CM,lab} \leq W_{dqo}$ and $W_{CM,field} \leq W_{CM,lab}$: the Candidate Method is accepted as equivalent to the Standard Method.

9 TEST PROGRAMME 3 – METHODS FOR PARTICULATE MATTER

9.1 General

This test programme describes a procedure for determining whether a Candidate Method (CM) is suitable to be considered equivalent to the EN Reference Method for the measurement of particulate matter in ambient air, using manual or automated measuring systems.
The EN Reference Methods have been developed to meet the Data Quality Objectives of the Daughter Directives of the Air Quality Framework Directive. The CM will also have to meet the directives' requirements, as explained earlier in the main document.

This test programme is suitable to evaluate CM for monitoring the PM_{10} or $PM_{2.5}$ fraction of total suspended particulates in ambient air. For example, this methodology may be used to evaluate alternative sample inlets, automated methods such as those based on the use of oscillating microbalances or ß-ray attenuation. Also other methods, such as in-situ optical methods may be evaluated for application below the upper assessment threshold specified in the relevant Daughter Directive.

9.2 Overview of the test procedure

The EN Standard Methods specify procedures to determine the expanded uncertainty of the method, e.g., by using a 'GUM' approach. This expanded uncertainty was compared with the expanded uncertainty requirement given in the relevant Daughter Directive, and was shown to fulfil this requirement.

Testing for equivalence will normally be carried out in two parts: a laboratory test in which the contributions of the different uncertainty sources to the measurement uncertainty will be assessed, and a field test in which the Candidate Method will be tested side-by-side with the relevant Standard Method.

For methods for particulate matter laboratory tests are only relevant if the CM is a modification of the existing EN Standard Method, in which case the field test will not be required. Generally, the test procedure will consist of a series of field tests in which the Candidate Method is tested side-by-side with the relevant Standard Method. In general, analysis of filter samples for manual methods will be performed by gravimetric measurement of the mass of particulates collected in conformity with the weighing procedures described in refs. [8] and [29].

When testing Candidate Methods based on the use of sample inlets differing from those applied in the EN Standard Methods, a more sensitive test for equivalence consists of the comparison of the filter contents of the soluble fraction of tracer ions that are suitable for the specific cut-off for PM_{10} or $PM_{2.5,}$ such as calcium, sodium or magnesium (PM_{10}) or sulphate, ammonium or nitrate ($PM_{2.5}$) [30].

The Candidate Method should be tested in a way that is representative for its practical use; frequencies of tests and re-calibrations used in practice should be applied in the Test Programme.

Results of existing studies, when demonstrably obtained according to the requirements of this test procedure, may be used to determine standard uncertainties. This is particulary relevant to the estimation of between-sampler/instrument uncertainties as described in 9.5. Results from prior studies may be used provided that they are obtained under conditions in accordance with the requirements of 9.4 and fulfill the criteria given in 9.5.

9.3 Laboratory test programme

9.3.1 General

The laboratory test programme is relevant for the following modifications of the Standard Method:

- Application of automated filter changers leading to filter storage conditions deviating from those prescribed in the EN Standard Methods
- Use of different weighing conditions, e.g., conditions deviating from the requirements set in the EN Standard Methods.

9.3.2 Application of automated filter changers

The assessment of the effect of applying automated filter changers shall be assessed as follows.

Worst-case conditions at monitoring sites shall be established. These must reflect the most unfavourable storage temperatures, using both average day and night time temperatures, for the maximum storage time, in situations when significant fractions of semi-volatile materials are expected on the filters. The storage temperature will depend on a combination of the ambient temperature and the effects of both insolation and local sources of heating and cooling. In general, worst-case effects will not be seen at times of continuously high ambient temperatures, but when storage temperatures are higher than those during sampling. In order to identify worst-case conditions both temperatures to which the sampled filters are exposed and ambient temperatures shall be measured.

A minimum of 40 samples shall be collected in conditions known to produce significant fractions of semi-volatile material on the filter.

These samples shall be removed from the sampler and weighed according to the procedure of the EN Standard Method. Subsequently, the samples shall be exposed to the worst-case conditions of time and temperature established, in a temperature-controlled cabinet, and reweighed according to the procedure of the EN Standard Method.

The largest mass loss observed shall be entered into the uncertainty budget as the 'loss due to storage' $u_{storage}$ by conversion assuming a uniform distribution:

$$u_{storage}^2 = \frac{(\Delta m)^2}{12} \qquad (9.1)$$

where
Δm = the largest mass loss observed for a single sample.

9.3.3 Different weighing conditions

The additional uncertainty arising from the use of weighing conditions outside the range specified in the EN Standard Method shall be assessed both for blank filters and for samples. For the latter, worst-case conditions of particulate composition shall be selected, by consideration of the mass of hygroscopic and semi-volatile materials sampled.

A minimum of 5 blank filters, from at least 2 different batches, for each type of filter to be used in the field, shall be investigated. The mass change of the filters between the extremes allowed by the revised conditions, i.e., the limits of high temperature and high relative humidity, and low temperature and low relative humidity, shall be established. The maximum mass change of the filter shall be entered into the uncertainty budget as the difference due to weighing conditions $u_{w,blank}$ by conversion assuming a uniform distribution:

$$u^2_{w,blank} = \frac{(\Delta m)^2}{12} \quad (9.2)$$

where
Δm = the largest mass loss observed for a single blank filter.

A minimum of 40 samples shall be collected in conditions known to produce significant effects on filter mass when weighed under the weighing conditions proposed.
These samples shall first be weighed under conditions fulfilling the requirements of the relevant EN Standard Method, and subsequently under the new weighing conditions proposed.

The largest mass difference observed shall be entered into the uncertainty budget as the difference due to weighing conditions $u_{w,sample}$ by conversion assuming a uniform distribution:

$$u^2_{w,sample} = \frac{(\Delta m)^2}{12} \quad (9.3)$$

where
Δm = the largest mass loss observed for a single sample.

9.4 Field test programme

9.4.1 General

Field tests shall be performed in which Candidate and Standard Method are compared side-by-side. The measurements will serve to assess

- 'between-sampler/instrument' uncertainty of the Candidate Method through the use of two samplers or instruments
- 'comparability' of the Candidate and Standard Methods.

Generally, results of existing studies, when demonstrably obtained according to the requirements of this test procedure, may be used to determine standard uncertainties. This is particulary relevant to the estimation of between-sampler/instrument uncertainties (see also 9.2).
In order to assure proper functioning of the Standard Method, two reference samplers or instruments may be used. In this case the mean squared difference of the results of both reference samplers/instruments can be used as an estimate of the (random) uncertainty of the Reference Method for these tests(see 9.6).

9.4.2 Experimental conditions

Test sites shall be representative for typical conditions for which equivalence will be claimed, including possible episodes of high concentrations. A minimum of 4 comparisons shall be performed with particular emphasis on the following variables, if appropriate:

- Composition of the PM fraction, notably high and low fractions of semi-volatile particles, to cover the maximum impact of losses of semi-volatiles
- Air humidity and temperature (high and low) to cover any conditioning losses of semi-volatiles during the sampling process
- Wind speed (high and low) to cover any dependency of inlet performance due to deviations from ideal behaviour as dictated by mechanical design, or deviations from the designated sampling flow rate.

A minimum of 40 measurement results each averaged over at least 24-hour per comparison shall be collected.

Samplers and instruments shall be positioned in such a way that the effect of spatial inhomogeneities of the compound concentration in the sampled air are negligible in comparison with other uncertainty contributions.

Both methods shall be operated under conditions reflecting practical application in the field, e.g., calibration intervals, flow checks, analysis of blank samples.

During the tests, the following information shall be collected and recorded
- calibration procedures, equipment and intervals
- (results of) quality checks
- Temperature and pressure of the sampled air
- Other conditions relevant for the measurements performed (e.g., air humidity)
- Particular events/situations that may be of influence on measurement results.

9.5 Evaluation of data collected

9.5.1 Suitability of datasets

Of the full dataset at least 20% of the results shall be greater than 50% of the limit value specified in the relevant Daughter Directive.
Data shall only be removed from the data set when sound technical reasons can be found for doing so. All valid data shall be used for further evaluation.

Note: Indications of outlying data (pairs) may be obtained using Grubb's tests on the individual single-period variances. Outlier tests are to be performed at the 99% level.

9.5.2 Calculation of performance characteristics

9.5.2.1 Between-sampler/instrument uncertainty

First, the Candidate Method results for each 24-hour measurement period i are averaged for each sampler/instrument to give 24-hour values y_i.

The between-sampler uncertainty u_{bs} is calculated from the differences of all 24-hour results of the candidate samplers/instruments operated in parallel as:

$$u_{bs}^2 = \frac{\sum_{i=1}^{n} (y_{i,1} - y_{i,2})^2}{2n} \qquad (9.4)$$

where
$y_{i,1}$ and $y_{i,2}$ are the results of parallel measurements for a single 24-hour period i
n = number of 24-hour measurement results.

The between-sampler/instrument uncertainty shall be determined
- for each of the comparisons separately
- for all results together
- for two datasets obtained by splitting the full dataset according to PM concentrations: greater than or equal to, or below 50% of the upper assessment threshold specified in the relevant Daughter Directive.

A between-sampler/instrument uncertainty > 3 $\mu g.m^{-3}$ is an indication of unsuitable performance of one or both samplers/instruments, and equivalence shall not be declared for the Candidate Method when the criterion of between sampler/instrument uncertainty of $\leq$ 3 $\mu g.m^{-3}$ is not satisfied.

9.5.2.2 Comparison with the Standard Method

First, the performance of the reference samplers/instruments is checked by calculation of the relative between-sampler/instrument uncertainty as in eq. (9.5). The between-sampler/instrument uncertainty for the Standard Method shall be $\leq 2\ \mu g.m^{-3}$.

For a comparison with the Standard Method the results of the parallel measurements of candidate samplers/instruments are averaged to give one result y_i for each 24-hour measurement period.

For the evaluation of the uncertainty due to the 'lack of comparability' between Candidate and Standard Methods it is assumed that the relationship between measurement results of both methods can be described by a linear relation of the form

$$y_i = a + bx_i \qquad (9.5)$$

Note: in practice, the actual relationship between measurement results of manual and automated methods may not always be linear.

The relation between the average results of the Candidate Method and the average results of the Standard Method is established using a regression technique that leads to a symmetrical treatment of both variables. A commonly applied technique is orthogonal regression [7,28]. Algorithms for the calculation of *a* and *b* and their variances are given in Annex B.

The regression equation is calculated for

- each of the comparisons separately
- for all results together
- for a dataset representing PM concentrations greater than or equal to 50% of the upper assessment threshold specified in the relevant Daughter Directive.

In each case, a precondition for acceptance of the dataset is that:

- the slope *b* is insignificantly different from 1: $|b\text{-}1| \leq 2.u(b)$, and
- the intercept *a* is insignificantly different from 0: $|a| \leq 2.u(a)$,

where u(b) and u(a) are the standard uncertainties of the slope and intercept, respectively, calculated as the square root of their variances.
If these preconditions are not met, corrections for slope and/or intercept may be applied (see Clause 9.7).

For further evaluation, the uncertainty in the results of the Candidate Method from comparison with the Standard Method, $u_{C\text{-}s}$, is calculated using a general equation describing $u_{C\text{-}s}$ as a function of PM concentration x_i. The use of one general equation avoids having to distinguish between 3 different situations when a and/or b differ from 0 and 1, respectively.

The general relationship describing the dependence of $u_{C\text{-}s}$ on x_i is given by

$$u^2_{C\_S}(y_i) = \frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)x_i]^2 \qquad (9.6)$$

where

RSS = the sum of (relative) residuals resulting from the orthogonal regression

$u(x_i)$ = random uncertainty of the reference method; as such, the value of u_{bs} calculated for the application of the Standard Method in these tests may be used. (See also note below Eq. 7.30)

RSS, the sum of (relative) residuals is calculated using eq. 9.7.

$$RSS = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \qquad (9.7)$$

The uncertainty u_{C_S} is calculated for

- each of the comparisons separately
- for all results together
- for a dataset extracted from all results representing PM concentrations greater than or equal to 50% of the upper assessment threshold specified in the relevant Daughter Directive.

9.5.3 Calculation of the combined uncertainty to be assigned to the Candidate Method[*]

For all 7 datasets the combined relative uncertainty of the Candidate Method $w_{c,CM}$ is calculated by combining the contributions found in 9.5.2.1 and 9.5.2.2 as follows:

$$w_{c,CM}^2(y_i) = \frac{u_{C\_S}^2(y_i)}{y_i^2} \qquad (9.8)$$

For each of the datasets the uncertainty at the daily limit value w_{CM} is calculated by taking as y_i the concentration at the limit value.

9.5.4 Calculation of the expanded uncertainty of Candidate Method

For each of the datasets the expanded relative uncertainty of the results of the Candidate Method is calculated by multiplying $w_{c,CM}$ by a coverage factor k reflecting the appropriate number of degrees of freedom resulting from the determination of $w_{c,CM}$ as

$$W_{CM} = k \cdot w_{CM} \qquad (9.9)$$

In view of the large number of experimental results available, a coverage factor k=2 can be used.

9.6 Evaluation of results of field tests

The highest resulting uncertainty estimate W_{CM} is compared with the expanded relative uncertainty based on the data quality objective for the Standard Method W_{dqo}.
In principle, two cases are possible

1. $W_{CM} \leq W_{dqo}$: the Candidate Method is accepted as equivalent to the Standard Method

2. $W_{CM} > W_{dqo}$: the Candidate Method is not accepted as Equivalent Method.

9.7 Application of correction factors or terms

When case 2 in Clause 9.6 occurs, it is permitted to apply a correction factor or term resulting from the regression equation obtained for the full dataset. The corrected values obtained shall satisfy the requirements for all datasets or subsets (see 9.5.2.2).
However, even when the Candidate Method passes the equivalence test (Case 1 in Clause 9.6) a correction may be applied in order to improve the accuracy of the Candidate Method.

With reference to Clause 9.5.2.2, three distinct situations may arise.

[*] Both eqs. (9.6) and (9.7) contain an uncertainty component due to the implementation of the Standard Method.

1. The slope b is not significantly different from 1: |b-1| ≤ 2.u(b), the intercept a is significantly different from 0: |a| > 2.u(a).

In this case, the value of intercept a may be used as a term used to correct all input values y_i as follows:

$$y_{i.corr} = y_i - a \tag{9.10}$$

The resulting values of $y_{i,corr}$ may then be used to calculate by linear regression (eq. 9.5) a new relationship to calculate

$$y_{i.corr} = c + dx_i \tag{9.11}$$

$u_{C\_R,corr}$ is then calculated as

$$u^2_{C\_S}(y_{,corri}) = \frac{RSS}{(n-2)} - u^2(x_i) + [c + (d-1)x_i]^2 + u^2(a) \tag{9.12}$$

where u(a) = the uncertainty of the original intercept a, the value of which has been used to obtain $y_{i,corr}$ (see Annex B for calculation of u(a)). RSS is calculated using eq. (9.13).

$$RSS = \sum_{i=1}^{n} (y_i - c - dx_i)^2 \tag{9.13}$$

2. The slope b is significantly different from 1: |b-1| > 2.u(b), the intercept a is not significantly different from 0: |a| ≤ 2.u(a).

In this case, the value of the slope *b* may be used as a factor to correct all input values y_i as follows:

$$y_{i.corr} = \frac{y_i}{b} \tag{9.14}$$

The resulting values of $y_{i,corr}$ may then be used to perform a new linear regression to calculate $u_{C\_S,corr}$ as

$$u^2_{C\_S}(y_{i,corr}) = \frac{RSS}{(n-2)} - u^2(x_i) + [c + (d-1)x_i]^2 + x_i^2 u^2(b) \tag{9.15}$$

where u(b) is the uncertainty of the original slope b, the value of which has been used to obtain $y_{i,corr}$ (see Annex B for calculation of u(b)). RSS is calculated using eq. (9.13).

3. The slope b is significantly different from 1: |b-1| > 2.u(b), AND the intercept a is significantly different from 0: |a| > 2.u(a).

In this case, the values of the slope b and the intercept a may be used to correct all input values y_i as follows:

$$y_{i.corr} = \frac{y_i - a}{b} \tag{9.16}$$

The resulting values of $y_{i,corr}$ may then be used to perform a new linear regression to calculate $u_{C\_S,corr}$ as

$$u^2_{C\_S}(y_{i,corr}) = \frac{RSS}{(n-2)} - u^2(x_i) + [c + (d-1)x_i]^2 + x_i^2 u^2(b) + u^2(a) \quad (9.17)$$

where u(b) is the uncertainty of the original slope b, the value of which has been used to obtain $y_{i,corr}$ (see Annex B for calculation of u(b)), and u(a) is the uncertainty of the original intercept a, the value of which has been used to obtain $y_{i,corr}$ (see Annex B for calculation of u(a)). RSS is calculated using eq. (9.13).

Note: eq. (9.17) is a simplification because it does not include covariance between slope and intercept. The resulting uncertainty may be higher than when a covariance term is included.

The resulting values for $u_{C\_S,corr}$ can then be entered in eq.(9.8) to calculate the combined relative uncertainty of the Candidate Method after correction as

$$w^2_{c,CM,corr} = \frac{u^2_{C\_S,corr}(y_i)}{y_i^2} \quad (9.18)$$

and the expanded relative uncertainty $W_{CM,corr}$ as

$$W_{CM,corr} = k \cdot w_{CM,corr} \quad (9.19)$$

$W_{CM,corr}$ can then be re-evaluated as in Clause 9.6.

9.8 Examples

In annex D some examples are given of results of equivalence testing for AMS for particulate matter according to the above procedures.

9.9 Ongoing QA/QC of the Equivalent Method

There is a requirement for ensuring the ongoing quality of the particulate measurement results obtained using the Equivalent Method. This is particularly important because the equivalence procedure depends on only field tests between the reference and Equivalent Method, and there is limited QA/QC that can be carried out on a routine basis (flow calibration). In addition, the equivalence tests were necessarily carried out under a limited range of particulate compositions, which may not continue to be representative for the actual conditions. Therefore, it may be necessary that periodic side-by-side comparisons are carried out between the reference and the Equivalent Methods at the sites where it is judged that the particulate composition is no longer representative of that during the equivalence tests.

10 TEST PROGRAMME 4 – SPECIATED PARTICULATE MATTER

10.1 General

This test programme describes a procedure for determining whether a Candidate Method (CM) is suitable to be considered equivalent to the EN Standard Method for the measurement of speciated substances in particulate matter in ambient air. The EN Standard Methods have been developed to meet the Data Quality Objectives of the Air Quality Framework Directive and its Daughter Directives. Therefore CM will similarly have to meet the directives' requirements, as explained earlier in the main document.

This test programme is suitable to evaluate CM for monitoring metals and PAH.

For example, this methodology may be used to evaluate the alternative analytical technique of inductive-coupled plasma – optical emission spectrometry for the measurement of metals or capillary electrophoresis for the measurement of benz[a]pyrene. Where only a small part of the method has been changed (variation on a theme such as a different extraction technique), then only the part of the method that is different needs to be investigated, by the laboratory tests detailed below.

10.2 Overview of the test procedures

The EN Standard Methods specify procedures to determine the expanded uncertainty of the method from the components performance criteria obtained during tests. This expanded uncertainty was compared with the expanded uncertainty criterion given in the relevant (proposed) Daughter Directive, and was shown to be within the required performance of the method with respect to the Directives' requirements.

Testing for equivalence will normally be carried out in two parts: a laboratory test in which the contributions of the different uncertainty sources to the measurement uncertainty will be assessed, and a field test in which the Candidate Method will be tested side-by-side with the relevant Standard Method.

If a CM is a modification to an existing EN Standard Method, then only the laboratory performance characteristics that are affected by the modification shall be tested and their standard uncertainties calculated. The standard uncertainties associated with the affected performance characteristics shall then be used together with these existing standard uncertainties for the other characteristics, to determine a new standard combined measurement uncertainty, $u_{c.}$

If a CM utilises a measurement method that is different to the Standard Method, then all of the tests shall be performed.

In both cases the results of existing studies, when demonstrably obtained according to the requirements of this test procedure, may be used to determine standard uncertainties.

The CM should be tested in a way that is representative for its practical use; frequencies of tests (e.g., response drift) and re-calibrations (e.g., flow rates) used in practice should be applied in the test programmes).

10.3 Laboratory Test Programme

10.3.1 General

In the laboratory test programme, the uncertainty sources listed in Table 3 are considered and assessed.

Table 3. Laboratory Test Programme 4: Uncertainty sources

Uncertainty source	**Symbol**
1 Mass of compound in sample	m_{sam}
1.2 Compound stability	*A*
1.3 Extraction/desorption efficiency	*D*
1.4 Mass of compound in calibration standards	m_{CS}
1.5 Response factors 1.5.1 lack-of-fit of calibration function 1.5.2 analytical repeatability 1.5.3 drift between calibrations	*F*
1.6 Selectivity	*R*
2 *Mass of compound in blank*	m_{bl}

The uncertainty sources that require assessment depend on the differences between Candidate and Reference Methods as follows:

Is the Candidate Method based on a different measurement principle?
In that case, the full TP needs to be performed.

Is the Candidate Method a modification of the Reference Method ?
In this case, the uncertainty sources relevant to the modification need to be investigated, e.g.

- 1.3 and 1.6 for alternative extraction solvents
- 1.5 and 1.6 for alternative analytical configurations.

10.3.2 Test programme

10.3.2.1 Mass of compound sampled

The mass of a compound sampled may be expressed as

$$m_{sam} = \frac{m_{meas}}{E \cdot A \cdot D} \qquad (10.1)$$

where
E = sampling efficiency
A = compound stability in the sample
D = extraction/desorption efficiency
m_{meas} = mass of compound measured in the analytical sample (extract, desorbate).

A correction for extraction/desorption efficiency is only applied when D is significantly different from 1 (see 10.3.2.1.3).

10.3.2.1.1 *Sampling efficiency*

For the purpose of this Test Programme the sampling efficiency is considered to be a part of the sampling procedure and, hence, is not dealt with. There may be problems, for example due to losses or degradation of compounds (e.g., benz[a]pyrene), but these will not affect the equivalence of the part(s) of the method under consideration in this Test Programme.

10.3.2.1.2 *Compound stability*

The compound stability shall be experimentally established for storage under conditions (time, temperature, environment) typical to the individual laboratory.

Tests shall be performed at a compound level corresponding to the ambient air limit or target value.

At times t=0 and t=t, n samples each shall be analyzed under repeatability conditions ($n \geq 6$). For both times the samples shall be randomly picked from a batch of representative samples in order to minimize possible systematic concentration differences. As a test of (in)stability a t-test will be performed (95% confidence, 2-sided). The t-test must show no significant difference between the start and end of the stability test.

The uncertainty of the stability determination consists of contributions from
- extraction/desorption (random part of extraction/desorption efficiency)
- calibration (random part of calibration)
- analytical precision
- inhomogeneity of the sample batch.

As such, the contribution of the determination of stability will already be incorporated in other contributions and needs not to be taken into account in the uncertainty.

10.3.2.1.3 *Extraction/desorption efficiency*

The extraction/desorption efficiency of the compound from the sample and its uncertainty are typically obtained from replicate measurements on Certified Reference Materials (CRM). For metals and benz[a]pyrene no CRM exist that are representative for the samples obtained; in the absence of such CRM, NIST SRM for total suspended particulates may be used to evaluate extraction efficiency. A minimum of 6 replicate measurements shall be performed.
The uncertainty due to incomplete extraction/desorption for the level corresponding to the limit value is calculated from contributions of
- the uncertainty in the concentration of the CRM
- the standard deviation of the mean mass determined

as

$$\frac{u^2(D)}{D^2} = \frac{u^2(m_{CRM}) + \frac{s^2(m_D)}{n}}{m_{CRM}^2} \qquad (10.2)$$

where
m_{CRM} = certified mass in the CRM
$s(m_D)$ = standard deviation of the replicate measurement results of the mass determined
n = the number of replicate measurements of the CRM.

When D is significantly different from 1 (at the 95% confidence level), the measurement result shall be corrected accordingly (see eq. (10.1)).

The value of $s(m_D)$ is used as an indicator of the relative uncertainty due to analytical repeatability w_{anal}:

$$w_{anal}^2 = \frac{s^2(m_D)}{m_D^2} \qquad (10.3)$$

10.3.2.1.4 *Measured mass of compound*

The uncertainty in the merasured mass of a compound determined by
- the uncertainty in the concentrations of the calibration standards used
- the lack-of-fit of the calibration function

- drift of detector response between calibrations
- the precision of the analysis
- the selectivity of the analytical system used.

Calibration standards

The calibration standards used will consist of solutions of the analyte; the uncertainty in the concentrations will be built up of contributions from

- the purity of the compound used; as the compounds under study are generally available in purities > 99%, the contribution of the purity may be considered insignificant
- when gravimetry is used to prepare the calibration solutions: the uncertainties in the weighings of compounds and solutions
- when volumetric techniques are used to prepare the calibration solutions: the uncertainties in the calibrated volumes of glassware and syringes used.

Note: Examples of calculations of uncertainties can be found in ref. [21].

Lack-of-fit of calibration function

The relative uncertainty due to lack-of-fit of the calibration function can be calculated for the relevant concentration (corresponding to the mass of benzene sampled at the limit value) from parameters obtained by a least-squares linear regression ($r = a + b.m_{cs}$), weighted in the concentration of the calibration standard.

Note: Options for the calculation of the uncertainty are given in ref. [21], Appendix E3 (equations E3.3 to E3.6).

As a worst-case approach, the relative uncertainty shall be estimated as

$$w_F^2 = \frac{u^2(m_r)}{m_r^2} = \frac{u^2(r) + s^2(a) + s^2(b)m_r^2}{b^2 m_r^2} \qquad (10.4)$$

where
m_r = mass calculated from the regression equation at response r
u(r) = uncertanty in the response r
b = slope of calibration function
a = intercept of calibration function
s = standard deviation of parameter between parentheses.

Response drift between calibrations

Normally, the current response factor will be used until a new one is established. In the interval between the re-establishment of its uncertainty, response checks – and, when necessary, adjustments of response factors - shall be performed as an element of ongoing quality control.
In the interval before the next checks response drift may occur. The relative uncertainty due to response drift for the period between subsequent adjustments of response factors shall then be estimated from the relative differences in responses between subsequent checks, as

$$w_d^2 = \frac{(r_n - r_{n-1})^2}{3\left(\frac{r_n + r_{n-1}}{2}\right)^2} \qquad (10.5)$$

where r_n is the detector response for a calibration standard corresponding closest to the mass representing a sample at the limit value. This approach assumes that no correction is applied for response drift, e.g., by averaging of subsequently determined response factors.

Selectivity

The analytical system used shall be optimized in order to minimize uncertainty due to the presence of potential interferents. Tests shall be performed with typical interferents at levels corresponding to 5 times the limit value of the compound under study. The uncertainty due to interferences may be obtained from ISO 14956 [23] as

$$w_R^2 = \frac{(r_+ - r_0)^2}{3r_0^2} \tag{10.6}$$

where r_+ represents the response with interferent, and r_0 represents the response without.

10.3.2.1.5 *Combined uncertainty in the sampled mass*

The contributions given above are combined to give the uncertainty of the mass of compound in the air sample as

$$\frac{u^2(m_{sam})}{m_{sam}^2} = \frac{u^2(m_{cs})}{nm_{cs}^2} + w_{anal}^2 + w_F^2 + w_d^2 + w_R^2 \tag{10.7}$$

where
n = number of calibration standards used to construct the calibration function (≥5)
w_R = relative uncertainty due to (lack of) selectivity of the analytical system.

10.3.2.2 Mass of compound in sample blank

The mass of compound in a sample blank is determined by analysis under repeatability conditions of a series of sample blanks; a minimum of 6 replicate analyses should be performed. The uncertainty is then calculated using the slope of the calibration function extrapolated to the blank response level as

$$u^2(m_{bl}) = \frac{s_{bl}^2}{nb_{bl}} \tag{10.8}$$

where
s_{bl} = standard deviation of the replicate blank analyses
n = number of replicate analyses
b_{bl} = slope of the calibration function at the blank response level.

When the blank response is below 3 times the noise level of the detector, then the blank level and its uncertainty shall be calculated from the detector noise level using the slope of the calibration function extrapolated to zero response assuming a uniform distribution as

$$m_{bl} = \frac{3r_0}{2b_0} \tag{10.9}$$

$$u^2(m_{bl}) = \frac{9r_0^2}{12} \tag{10.10}$$

where
r_0 = noise level
b_0 = slope of calibration function at zero response.

10.3.2.3 Combined uncertainty

The combined relative uncertainty of the compound mass in the air sampled is obtained by combination of contributions given in Clauses 10.3.2.1 – 10.3.2.2 as

$$w^2_{CM,lab} = \frac{u^2_{lab}(m)}{m^2} = \frac{u^2(m_{sam}) + u^2(m_{bl})}{(m_{sam} - m_{bl})^2} \qquad (10.11)$$

10.2.3.4 Expanded uncertainty

The expanded relative uncertainty of the Candidate Method resulting from the laboratory experiments, $W_{CM,lab}$ at the 95% confidence level is obtained by multiplying $w_{CM,lab}$ with a coverage factor appropriate to the number of degrees of freedom resulting from the performance of the test programme. This can be calculated by applying the Welch-Satterswaithe equation (ISO-GUM, H2). For a large number of degrees of freedom, a coverage factor of 2 is used.

Note: as a first approximation, the number of degrees of freedom may be based on that of an uncertainty contribution covering more than 50% of the variance budget.

10.3.2.5 Evaluation of results of the laboratory tests

The resulting $W_{CM,lab}$ is compared with the expanded relative uncertainty based on the data quality objective for the Standard Method W_{dqo}.
If $W_{CM,lab} \leq W_{dqo}$, the field test programme can be performed; if not, the Candidate Method shall first be improved, and relevant changes tested in the laboratory test programme.

10.4 Field test programme

10.4.1 General

When required, field tests shall be performed in which Candidate and Standard Method are compared side-by-side. The measurements will serve to assess

- 'between-sample' uncertainty of the Candidate Method through the use of replicate samples
- 'comparability' of the Candidate and Standard Methods.

For constituents of particulate matter, the sampling is not a part of the equivalence testing. Therefore, sub-samples from high-volume samples with different loadings may be used to obtain the required information. In principle, 8 or more sub-samples may be obtained from one high-volume sample and the homogeneity of compound loadings on the sub-samples has been demonstrated for benz[a]pyrene [31] to be better than < 4 % (coefficient of variation) when applying the Standard Method.

In order to assure proper functioning of the Standard Method, a minimum of two samples shall be analyzed by application of the Standard Method.

The number of replicate samples needed to determine the between-sampler uncertainty of the Candidate Method (Standard Method) will depend on whether the Candidate Method is to be used by more than one laboratory. When used by one laboratory, a minimum of six sub-samples will be analyzed using the Candidate Method.

When used by more than one laboratory, the field test is also used to assess between-laboratory contributions to the uncertainty of the measurement results. For this purpose, each laboratory will analyze a minimum of two samples using the Candidate Method.

10.4.2 Experimental conditions

Samples shall be representative of typical conditions for which equivalence will be claimed, including possible episodes of high concentrations. A minimum of 4 comparisons shall be performed with particular emphasis on the following variables, if appropriate:

- Composition of the air, notably high and low concentrations of the measured compound and potential interferents
- Air humidity and temperature (high and low) to cover any effects on extraction efficiency.

For the Candidate Method a minimum of 20 different high-volume samples per comparison – to be divided into 8 sub-samples each - shall be collected. Alternatively, a minimum of 160 samples obtained using a low-volume reference sampler may be used.

Samplers and instruments shall be positioned in such a way that the effect of spatial inhomogeneities of the compound concentration in the sampled air are negligible in comparison with other uncertainty contributions.

Both methods shall be operated under conditions reflecting practical application in the field, e.g., calibration intervals, response checks, analysis of blank samples.

During the tests, the following information shall be collected and recorded

- Calibration procedures, equipment and intervals
- (Results of) quality checks
- Other conditions relevant for the analyses performed.

10.4.3 Evaluation of test results

10.4.3.1 Suitability of the dataset

Of the full dataset, at least 20% of the results shall be greater than or equal to the upper assessment threshold specified in the relevant (currently proposed) Daughter Directive.
Data shall only be removed from the data set when sound technical reasons can be found for doing so. All valid data shall be used for further evaluation.

Note: Indications of outlying data within replicate sets may be obtained using Grubb's tests on the individual single-period variances. Tests are to be performed at the 99% level.

10.4.3.2 Calculation of performance characteristics

10.4.3.2.1 Between-sampler/instrument uncertainty

The relative between-sample uncertainty for individual laboratories w_{bs} is calculated for the full dataset from the differences of results of the candidate samplers/instruments operated in parallel as:

$$w_{bs}^2 = \frac{\sum_{i=1}^{n} \left(y_{i,1} - y_{i,2}\right)^2}{2n\bar{y}^2} \quad \text{for duplicate samples} \qquad (10.12)$$

where
$y_{i,1}$ and $y_{i,2}$ are the results of parallel measurements for a single period *i*
$\bar{y}$ = average of all measurement results of the Candidate Method
n = number of measurement results.

$$w_{bs}^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\left(y_{i,j} - \overline{y_i}\right)^2}{n(p-1)\overline{y}^2} \quad \text{for replicate samplers with p > 2} \qquad (10.13)$$

where
y_{ij} = result of measurement *j* for a single period *i*
$\overline{y}_i$ = mean result for period i
p = number of replicates for period i.

When more than one analytical laboratory is participating, equation 10.13 shall be used to calculate the between-laboratory w_{bs}.

The w_{bs} between sampler uncertainty component for each individual laboratory and the between-laboratory w_{bs} (if relevant) shall comply with the criteria given in Annex A.

If the performance of a single laboratory causes a method implemented by more than two laboratories to fail the criteria, then the results for this laboratory may be excluded, if sound technical grounds exist for doing so.

10.4.3.2.2 *Comparison with Standard Method*

First, the performance of the reference samplers/instruments is checked by calculation of the relative between-sampler/instrument uncertainty as in eq. (10.12) or (10.13). The relative between-sampler/instrument uncertainty for the Standard Method shall be ≤ 4%.

For a comparison with the Standard Method first the results of replicate measurements are averaged to give data pairs 'Candidate Method – Standard Method' with equal measurement periods.

For the evaluation of the uncertainty due to the 'lack of comparability' between Candidate and Standard Method it is assumed that the relationship between measurement results of both methods can be described by a linear relation of the form:

$$y_i = a + bx_i \qquad (10.14)$$

where x_i is the average result of the Standard Method over period *i.*

The relation between the average results of the Candidate Method and those of the Standard Method is established for the full dataset using a regression technique that leads to a symmetrical treatment of both variables. A commonly applied technique is orthogonal regression [7,28].

The uncertainty due to lack of comparability will be a function of the concentration of the measurand.
The general relationship describing the dependence of $u_{C\text{-}S}$ on x_i is given by

$$u_{C\_S}^2(y_i) = \frac{RSS}{(n-2)} - u^2(x_i) + \left[a + (b-1)x_i\right]^2 \qquad (10.15)$$

where
RSS = the sum of (relative) residuals resulting from the orthogonal regression

$u(x_i)$ = random uncertainty of the Standard Method; as such, the value of u_{bs} calculated for the application of the Standard Method in these tests may be used. (See also note below Eq.7.30)

Algorithms for the calculation of *a* and *b* and their variances are given in Annex B.

RSS, the sum of (relative) residuals is calculated using eq. 10.16a or 10.16b, depending on whether the residuals or relative residuals are constant.

$$RSS = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \text{ when } (y_i - a - bx_i)^2 \text{ is constant} \qquad (10.16a)$$

$$RSS = (a + bx_i)^2 \sum_{i=1}^{n} \left(\frac{y_i}{a + bx_i} - 1 \right)^2 \text{ when } \left(\frac{y_i}{a + bx_i} - 1 \right)^2 \text{ is constant} \qquad (10.16b)$$

10.4.3.3 Calculation of the combined uncertainty of Candidate Method

The combined relative uncertainty of the Candidate Method $w_{c,CM}$ is calculated by combining the contributions found in 10.4.3.2.1 and 10.4.3.2.2 as follows:

$$w_{c,CM}^2(y_i) = w_{bs}^2 + \frac{u_{C\_S}^2(y_i)}{y_i^2} \qquad (10.17)$$

In this way, $w_{c,CM}$ is expressed as a function of the compound concentration.

The uncertainty at the limit value w_{CM} is calculated by taking as y_i the concentration at the limit value.

10.4.3.4 Calculation of the expanded uncertainty of Candidate Method

The expanded relative uncertainty of the results of the Candidate Method is calculated by multiplying $w_{c,CM}$ by a coverage factor k reflecting the appropriate number of degrees of freedom resulting from the determination of $w_{c,CM}$ as

$$W_{CM,field} = k \cdot w_{c,CM} \qquad (10.18)$$

In view of the large number of experimental results available, a coverage factor k=2 can be used.

10.4.4 Evaluation of results of field tests

The resulting uncertainty estimate $W_{CM,field}$ is compared with the expanded relative uncertainty obtained from the laboratory test programme $W_{CM,lab}$ and the expanded relative uncertainty based on the data quality objective for the Standard Method W_{dqo}.
In principle, three cases are possible

1. $W_{CM,field} \leq W_{CM,lab}$: the Candidate Method is accepted as equivalent to the Standard Method

2. $W_{CM,lab} < W_{CM,field} \leq W_{dqo}$: the Candidate Method is accepted conditionally; before final acceptance, the uncertainty evaluation from the laboratory tests should be revisited and corrected such that situation 1 occurs

3. $W_{CM,field} > W_{dqo}$: the Candidate Method is not accepted as Equivalent Method.

11 REPORTING REQUIREMENTS

Final reports on the Demonstration of Equivalence submitted to the Competent Authority and further to the European Commission should contain – at minimum – the following information.

Title of the method

Executive summary

General information

1. A summary of the principles of the Candidate Method; the full Standard Operating Procedure of the method, including a description of ongoing QA/QC, shall be annexed.
2. The scope of equivalence testing, i.e., the differences between the Candidate Method and the Reference Method that require specific tests to be performed.
3. A description of the conditions for which equivalence with the Reference Method is claimed, e.g., concentration range, environmental conditions, type of location.
4. Sources of uncertainty data for unchanged parts of the EN standards enacting the Reference Method, where relevant.
5. Names of the laboratories involved in the test programme(s) and the scope of their relevant competences, e.g., ISO 17025 accreditation.

Laboratory test programme (where applicable)

6. The parameters tested in the laboratory programme.
7. A description of the test procedures used, including procedures for the establishment and maintenance of measurement traceability where relevant, and procedures for quality control and quality assurance.
8. The test results, the results of the uncertainty assessment, and the results of their comparison with the relevant Data Quality Objectives including uncertainty or, in the absence of Data Quality Objectives, the results of the comparison between Candidate Method and Reference Method.

Field test programme (where applicable)

9. Full description of the test locations, test periods and conditions (e.g. temperature, humidity, wind velocity, concentration level)
10. A description of the equipment and test procedures used, including procedures for the establishment and maintenance of measurement traceability where relevant, and procedures for quality control and quality assurance.
11. The test results, the results of the uncertainty assessment, and the results of their comparison with the relevant Data Quality Objectives including uncertainty, or, in the absence of Data Quality Objectives, the results of the comparison between Candidate Method and Reference Method.

Conclusions

12. Results of the overall testing of the performance of the Candidate Method as compared to the Data Quality Objectives specified in the relevant EU ‘Daughter Directive’
13. The overall conclusion about the equivalence including restrictions, if any, in the conditions under which the claim to equivalence is valid or generalizations of the equivalence claim to other relevant conditions. Relevant conditions include concentration ranges, meteorological conditions, geographical locations and/or type(s) of monitoring sites.

12 REFERENCES

1. Council Directive 96/62/EC on ambient air quality assessment and management.

2. Council Directive 1999/30/EC relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air.

3. Council Directive 2000/69/EC relating to limit values for benzene and carbon monoxide in ambient air.

4. Council Directive 2002/3/EC relating to ozone in ambient air.

5. Council Directive 2004/107/EC relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air.

6. NEN 7778. Environment – Equivalency of measurement methods. NEN, Delft, 2002.

7. DIN 38402. Gleichwertigkeit von Analysenverfahren auf Grund des Vergleiches von Analysenergebnissen und deren statistischer Auswertung – Vorgehensweise für quantitative Merkmale mit kontinuierlichem Wertespektrum. DIN, Berlin, 2004.

8. EN 12341. Air quality – Determination of the PM_{10} fraction of suspended particulate matter – reference method and field test procedure to demonstrate reference equivalence of measurement methods. CEN, Brussels, 1998

9. EN 14211. Ambient air quality – Measurement method for the determination of the concentration of nitrogen dioxide and nitrogen monoxide by chemiluminescence. CEN, Brussels, 2004.

10. EN 14212. Ambient air quality – Measurement method for the determination of the concentration of sulphur dioxide by UV fluorescence. CEN, Brussels, 2004.

11. EN 14625. Ambient air quality – Measurement method for the determination of the concentration of ozone by UV photometry. CEN, Brussels, 2004.

12. EN 14626. Ambient air quality – Measurement method for the determination of the concentration of carbon monoxide by non-dispersive infrared spectrometry. CEN, Brussels, 2004.

13. EN 14662-1. Standard method for the determination of benzene in ambient air – Part 1: Method with pumped sampling, thermal desorption and capillary gas chromatography. CEN, Brussels, 2004.

14. EN 14662-2. Standard method for the determination of benzene in ambient air – Part 2: Method with pumped sampling, solvent desorption and capillary gas chromatography. CEN, Brussels, 2004.

15. EN 14662-3. Standard method for the determination of benzene in ambient air – Part 3: Method with automated gas chromatographs. CEN, Brussels, 2004.

16. EN 14662-4. Standard method for the determination of benzene in ambient air – Part 4: Method with diffusive sampling, thermal desorption and capillary gas chromatography. CEN, Brussels, 2004.

17. EN 14662-5. Standard method for the determination of benzene in ambient air – Part 5: Method with diffusive sampling, solvent desorption and capillary gas chromatography. CEN, Brussels, 2004.

18. EN 14902. Ambient air quality – Standard method for the measurement of Pb, Cd, As and Ni in the PM10 fraction of suspended particulate matter. CEN, Brussels, 2004.

19. ISO 6144. Gas analysis – Preparation of calibration gas mixtures – Static volumetric method. ISO, 2002.

20. ISO 6145. Gas analysis – Preparation of calibration gas mixtures using dynamic volumetric methods – all parts.

21. Eurachem / Citac Guide G4. Quantifying Uncertainty in Analytical Measurement, 2nd edition (2000).

22. Draft-VDI 2100, part 4. Gaseous ambient air measurement. Indoor air pollution measurement – Gas chromatographic determination of organic compounds – Calibration procedures as a measure for quality assurance. VDI, Düsseldorf, 2002.

23. EN-ISO 14956. Air quality – Evaluation of the suitability of a measurement method by comparison with a stated measurement uncertainty. ISO, Geneva, 2001.

24. EN 13528-1. Ambient air quality – Diffusive samplers for the determination of gases and vapours – Requirements and test methods – Part 1: General requirements.

25. EN 13528-2. Ambient air quality – Diffusive samplers for the determination of gases and vapours – Requirements and test methods – Part 2: Specific requirements and test methods.

26. EN 13528-3. Ambient air quality – Diffusive samplers for the determination of gases and vapours – Part 3: Guide to selection, use and maintenance.

27. Hafkenscheid, Th. L. Diffusive sampler validation using measurement uncertainty as performance characteristic. Proceedings of the International Conference Measuring Air Pollutants by Diffusive sampling. Montpellier, 2001. EUR 20242 EN, EC, 2002.

28. Massart, D.L.; A. Dijkstra, L. Kaufman. Evaluation and Optimization of Laboratory Methods and Analytical Procedures. Elsevier, Amsterdam, 1978.

29. prEN 14907. Ambient Air Quality – Reference gravimetric measurement method for the determination of the PM2.5 mass fraction of suspended particulate matter in ambient air.

30. Allegrini, I.; A. Febo, C. Perrino. Critical aspects of harmonization in particulate matter measurements: evaluation of methods and equivalence procedures. Paper presented at the International Conference on QA/QC in the field of emission and air quality measurements. Prague, 2003.

31. CEN TC 264 WG 21. Test protocol for field validation of B[a]P reference measurement methods. Working Group Document N67.

ANNEX A

Table A1. Criteria for between-sampler/instrument and between-laboratory uncertainties for specified compounds

Compound	*Required Standard Uncertainty (%) **	*Between-sampler/ instrument (%)*	*Between lab (%)*
Sulphur dioxide	7.5	5	5
Nitrogen dioxide	7.5	5	5
Ozone	7.5	5	5
Carbon monoxide	7.5	5	5
Benzene	12.5	3	7.5
Benz[a]pyrene	25	4	15
Nickel	20	5	12.5
Cadmium	20	5	12.5
Lead	12.5	4	7.5
Arsenic	20	5	12.5

* 50% of the data-quality objective expanded uncertainty for continuous or fixed measurements as specified in relevant (proposed) Daughter Directives.

ANNEX B

Algorithms for the calculation of orthogonal regression parameters.

Regression equation: y = a + b.x

Slope b:

$$b = \frac{Syy - Sxx + \sqrt{(Syy - Sxx)^2 + 4(Sxy)^2}}{2Sxy}$$

Where:

$$Sxx = \sum (x_i - \bar{x})^2$$

$$Syy = \sum (y_i - \bar{y})^2$$

$$Sxy = \sum (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

$$\bar{x} = 1/n \sum x_i$$

$$\bar{y} = 1/n \sum y_i$$

Intercept a:

$$a = \bar{y} - b \cdot \bar{x}$$

The uncertanties of the slope and intercept (for corrections to PM Candidate Methods):

$$u^2(b) = \frac{Syy - ((Sxy)^2 / Sxx)}{(n-2).Sxx}$$

$$u^2(a) = u^2(b) \frac{\sum x^2}{n}$$

ANNEX C

Members of the Equivalence Working Group (alphabetical order)

Pascual Perez Ballesta	European Commission – Joint Research Centre – Institute of Environment and Sustainability
Antonio Febo	Centro Nationale di Ricerca (Italy)
Rosalia Fernandez-Patier	Instituto Sanitad Carlos III (Spain)
Marina Fröhlich	Umweltbundesamt (Austria)
Saul dos Santos Garcia	Instituto Sanitad Carlos III (Spain)
Theo Hafkenscheid (chairman)	Nederlands Meetinstituut (Netherlands)
Stefan Jacobi	European Commission – Directorate-General Environment
Ton van der Meulen	Rijksinstituut voor Volksgezondheid en Milieu (Netherlands)
Don Munns	CEN Technical Committee 264 'Air Quality'
Hans-Ulrich Pfeffer	Landesumweltamt Nordrhein-Westfalen (Germany)
Jean Poulleau	INERIS (France)
Kevin Saunders	KERIS Ltd. (United Kingdom)
Jari Walden	Finnish Meteorological Institute (Finland)
Peter Woods	National Physical Laboratory (United Kingdom)

ANNEX D

EXAMPLES OF RESULTS OF EQUIVALENCE TESTING FOR AMS FOR PARTICULATE MATTER

Introduction

This annex contains some examples of results of equivalence testing for automated methods for measurement of $PM_{2.5}$, and PM_{10}, performed using the algorithms given in Clause 9.5-9.7. In these assessments, a limit value for $PM_{2.5}$ of 35 $\mu g.m^{-3}$ has been assumed following a recommendation in the Second Position paper on Particulate Matter.

The measurement data used for $PM_{2.5}$ were taken from the validation study carried out by CEN TC 264 WG 15. This annex gives three examples:

- One in which the CM for one location fulfills the requirements for equivalence but may still be improved in accuracy by correction of its results
- One in which the CM for one location does not fulfill the requirements until after correction
- On in which correction of results for one location improve the comparability, but does not lead to acceptance of the CM as an Equivalent Method.

The measurement data used for PM_{10} were supplied by the Landesumweltamt Nordrhein-Westfalen – LUA-NRW (Germany).

Using the algorithms presented in Clause 9.5 it is relatively easy to judge whether corrections may lead to an improvement beyond meeting the equivalence requirements: unless the slope *b* of the regression equation obtained for uncorrected results is high, a random uncertainty above 12,5% of the limit value concentration (about 4,4 $\mu g.m^{-3}$ for $PM_{2.5}$; about 6,3 $\mu g.m^{-3}$ for PM_{10}) is an indication that corrections will generally fail to bring the required improvement because of excessive scatter of the results of the reference and candidate methods when compared using regression, unless the slope of the regression equation is considerably higher than 1.

It should be noted that for $PM_{2.5}$ the between-sampler uncertainty for the reference method has been calculated from the data actually available from the validation study.
For PM_{10} a between-sampler uncertainty for the reference method of 1,5 $\mu g.m^{-3}$ has been assumed, based on information supplied by LUA-NRW.

Examples of results – $PM_{2.5}$

Comparison of CM1 and RM1 – location A

Full dataset (µg.m^{-3}, uncorrected)

RM1,1	*RM1,2*	*CM1*	*RM1,1*	*RM1,2*	*CM1*	*RM1,1*	*RM1,2*	*CM1*	*RM1,1*	*RM1,2*	*CM1*
20.2	21.4	18.4	18.8	22.0	24.9	15.3	15.4	19.6	19.9	20.5	23.7
15.1	16.5	15.9	17.7	18.4	20.8	22.3	21.7	26.4	22.3	24.2	26.0
17.8	16.8	20.0	17.6	19.0	18.2	22.7	24.2	26.4	25.4	25.8	25.1
20.1	19.3	21.8	23.4	23.5	21.8	22.3	22.5	26.4	20.4	21.5	21.2
24.6	26.7	29.2	23.5	24.3	24.0	19.7	21.3	25.0	18.4	17.8	20.2
30.3	29.9	37.0	17.0	17.0	21.2	17.3	17.9	23.5	19.4	19.8	22.8
24.0	24.2	27.7	16.9	17.1	19.5	23.2		26.0	35.9	37.3	42.2
21.2	21.4	23.4	26.4	27.7	37.9	24.4	25.2	27.2	23.5	24.6	29.2
19.4	20.8	23.3	27.4	28.6	29.5	17.6	17.2	22.4	25.8	26.0	30.8
22.3	21.9	22.7	34.0	36.1	31.8	10.0	10.8	14.9	12.4	12.9	18.0
14.9	15.5	17.0	25.7	26.2	28.1	21.5	26.8	26.7		11.8	15.6
13.2	12.5	15.7	29.3	31.1	31.8	13.5	13.8	12.7	10.7	11.2	15.8
11.0	12.4	12.7	19.2	19.9	22.9	7.8	8.1	9.8	12.6	13.2	17.4
12.1	13.0	14.6	28.2	30.6	28.9	13.5	13.2	16.6			
15.8	16.5	20.5	10.9	11.9	16.0	20.6	21.2	24.0			

Evaluation of uncorrected data for CM1

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	0.96	random term	1.89	µg/m³
uncertainty of b	0.05	bias at LV	-3.32	µg/m³
intercept a	-1.83	combined uncertainty	3.82	µg/m³
uncertainty of a	1.13	relative uncertainty	10.9%	pass
number of datapairs	58	RM between-sampler uncertainty	1.35	µg/m³

Evaluation of data for CM1 after correction for intercept

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	0.96	random term	2.20	µg/m³
uncertainty of b	0.05	bias at LV	-1.50	µg/m³
intercept a	0.00	combined uncertainty	2.66	µg/m³
uncertainty of a	1.13	relative uncertainty	7.6%	pass
number of datapairs	58	RM between-sampler uncertainty	1.35	µg/m³

Comparison of CM2 and RM1 – location B

Full dataset (µg.m^{-3}, uncorrected)

RM1,1	*RM1,2*	*CM2*	*RM1,1*	*RM1,2*	*CM2*	*RM1,1*	*RM1,2*	*CM2*	*RM1,1*	*RM1,2*	*CM2*
18.5	17.5	15.7	11.1	11.2	6.2	48.9	48.5	44.0	56.0	56.0	51.2
24.5	25.5	18.4	12.3	12.3	8.9	40.9	40.0	34.4	18.4	19.5	15.3
43.8	43.1	41.5	27.7	27.4	21.4	47.1	47.1	42.4	62.7	62.3	55.7
45.4	45.1	45.3	40.6	41.6	34.4	32.1	33.1	27.3	58.2	59.4	54.0
40.3	40.1	35.9	16.6	16.1	13.6	41.0	41.0	37.7	50.7	48.9	47.2
34.4	32.2	25.7	17.7	16.5	11.1	37.0	36.8	33.3	6.9	6.8	4.0
39.4	38.5	33.1	38.9	38.3	34.0	66.3	67.2	61.4	16.9	16.8	13.1
46.0	46.0	39.0	31.4	30.5	25.4	58.3	56.6	53.3	46.2	45.6	42.4
69.5	68.3	65.1	42.6	41.7	36.0	33.0	31.7	28.3	38.6	35.9	34.5
55.2	54.9	49.9	55.1	53.9	52.3	22.4	22.1	18.6	69.7	69.9	67.8
59.5	59.6	62.0	24.6	24.7	20.0	59.5	59.2	55.3	69.0	68.5	62.7
65.4	67.5	60.1	31.8	31.9	30.8	64.2	63.1	61.1	70.8	70.8	64.3
58.4	58.4	51.2	22.6	23.3	17.7	46.7	45.2	42.1	25.2	25.3	21.2
48.1	48.3	47.2	6.2	6.9	1.6	25.4	25.2	19.5	6.2	6.9	1.8
28.1	31.0	22.2	6.8	7.4	3.2	31.1	30.5	26.3	13.9	13.4	6.6
25.6	24.7	25.9	14.5	15.5	10.0	38.8	38.0	33.3	13.5	13.6	11.0
19.5	19.0	13.3	39.2	38.8	34.7	38.2	39.2	33.7	27.0	25.6	21.3
43.3	42.3	36.3	46.9	47.8	41.1	45.7	45.3	43.4	39.7	39.7	34.0
42.7	42.1	37.0	34.3	33.6	27.4	73.7	72.4	69.8	12.5	12.1	5.7
29.0	28.7	23.3	15.6	14.9	9.0	55.4	55.8	52.6	17.0	16.5	12.1
30.7	29.5	25.1	35.0	34.1	30.7	50.0	50.0	46.3	9.9	9.9	7.1
13.0	12.3	6.9	54.4	54.3	49.5	54.1	52.5	46.5			

Evaluation of uncorrected data for CM2

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	1.01	random term	1.29	µg/m³
uncertainty of b	0.01	bias at LV	-4.51	µg/m³
intercept a	-4.91 *	combined uncertainty	4.69	µg/m³
uncertainty of a	0.46	relative uncertainty	13.4%	fail
number of datapairs	87	RM between-sampler uncertainty	1.35	µg/m³

* Significant at 95% confidence level.

Evaluation of data for CM2 after correction for intercept

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	1.01	random term	1.38	µg/m³
uncertainty of b	0.01	bias at LV	0.40	µg/m³
intercept a	0.00	combined uncertainty	1.43	µg/m³
uncertainty of a	0.46	relative uncertainty	4.1%	pass
number of datapairs	87	RM between-sampler uncertainty	1.35	µg/m³

Comparison of CM1 and RM1 – location C

Full dataset (µg.m^{-3}, uncorrected)

RM1,1	*RM1,2*	*CM1*	*RM1,1*	*RM1,2*	*CM1*	*RM1,1*	*RM1,2*	*CM1*	*RM1,1*	*RM1,2*	*CM1*
14.8	15.1	12.4	33.0	34.8	16.1	16.9	17.0	13.2	8.4	8.2	7.7
4.8	6.2	4.5	23.2	24.4	12.1	10.0	10.1	7.7	10.9	10.8	9.7
19.1	19.8	20.4	20.7	22.6	11.1	7.4	7.6	8.2	14.8	14.8	9.9
13.5	16.1	11.4	14.4	15.4	9.1	12.0	11.8	8.5	38.5	37.3	24.5
18.8	19.9	13.8	8.1	8.6	5.9	10.9	11.0	8.7	40.0	41.1	23.6
22.2	21.5	12.8	7.1	7.2	9.9	22.3	22.6	13.4	17.0	17.8	11.6
7.2	8.2	4.0	10.7	11.3	8.9	17.8	17.8	7.4	16.0	15.5	14.8
16.4	16.0	11.2	13.7	14.1	12.3	8.8	11.2	2.9	27.1	26.5	18.4
43.7	46.1	39.9	22.7	22.8	15.9	5.6	5.7	7.5	25.7	27.2	20.3
37.5	40.2	26.1	12.5	12.6	11.3	10.8	10.0	8.9	35.8	35.8	17.3
48.2	52.0	34.6	8.4	7.3	7.6	6.3	7.2	6.4	13.8	14.1	5.2
44.3	47.4	31.3	6.3	6.2	6.8	6.2	8.0	5.5	18.6	18.6	12.5
25.8	26.9	16.0	11.6	11.6	11.8	6.9	6.9	7.3	33.8	34.1	20.3

Evaluation of uncorrected data for CM1

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	0.64 *	random term	2.77	µg/m³
uncertainty of b	0.04	bias at LV	-11.50	µg/m³
intercept a	1.21	combined uncertainty	11.82	µg/m³
uncertainty of a	0.80	relative uncertainty	33.8%	fail
number of datapairs	52	RM between-sampler uncertainty	1.35	µg/m³

* Significant at 95% confidence level.

Evaluation of data for CM1 after correction for slope

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	1.04	random term	4.89	µg/m³
uncertainty of b	0.06	bias at LV	2.51	µg/m³
intercept a	1.19	combined uncertainty	5.50	µg/m³
uncertainty of a	1.26	relative uncertainty	15.7%	fail
number of datapairs	52	RM between-sampler uncertainty	1.35	µg/m³

Examples of results – PM_{10}

Comparison of CM1 and RM – location D

Full dataset (µg.m^{-3}, uncorrected)

RM	*CM1*	*RM*	*CM1*	*RM*	*CM1*	*RM*	*CM1*	*RM*	*CM1*	*RM*	*CM1*
36.0	35.9	81.2	58.6	22.0	14.7	24.9	23.0	45.5	32.4	24.0	21.8
13.9	15.0	41.8	24.9	21.5	16.7	25.8	21.3	20.9	17.5	39.4	32.0
33.3	24.6	52.5	34.2	19.4	18.2	27.2	25.4	49.5	38.1	50.4	36.7
44.8	31.0	29.7	13.7	21.5	20.3	15.3	13.3	46.2	39.6	22.8	16.9
53.6	43.5	17.4	9.2	25.7	23.1	26.4	22.0	26.6	26.7	30.9	20.9
47.9	30.0	30.2	20.3	30.3	29.5	21.9	19.0	34.0	29.3	24.6	21.5
33.9	19.0	35.5	21.6	24.8	22.3	36.9	35.8	23.3	19.2	33.5	28.1
23.6	14.5	36.8	23.0	26.9	27.9	31.6	24.5	25.3	20.5	33.8	30.6
29.2	18.5	39.3	25.4	29.9	27.1	41.1	30.8	36.0	29.6	34.6	31.5
19.0	11.2	40.9	35.0	27.7	25.1	46.7	40.0	27.5	16.1	38.5	27.4
14.3	11.9	23.8	20.8	15.8	11.4	57.0	53.3	15.8	14.2	37.3	19.6
37.8	27.7	31.2	30.4	26.5	22.5	55.1	52.4	15.9	16.2	66.0	60.8
23.8	21.8	50.2	38.5	21.4	17.0	25.2	21.5	23.5	22.0	38.4	31.5
19.6	15.0	44.1	38.3	18.4	12.9	21.8	22.3	17.4	16.1	23.9	18.3
62.5	49.7	24.0	17.1	29.4	27.0	29.6	22.2	33.8	27.2	46.9	41.4
66.2	52.1	22.4	18.4	19.2	8.6	44.7	39.1	30.4	23.4	39.9	31.1
42.6	31.4	20.5	18.9	31.3	22.5	25.0	22.7	57.6	43.3	8.2	7.7
40.0	29.1	18.0	15.8	34.0	24.0	22.4	19.3	37.0	29.4	22.0	21.7
44.2	38.5	16.2	14.4	26.0	19.4	26.1	20.1	39.6	33.0	41.8	34.8
50.4	42.5	29.2	23.1	33.6	28.0	30.3	18.9	25.5	20.2	38.2	33.6
44.5	34.3	40.8	30.2	51.8	42.7	13.9	12.0	39.8	29.8	14.4	13.7
23.4	16.9	39.9	22.5	22.6	16.9	23.5	21.2	28.2	20.6	27.7	20.4
29.5	25.6	20.8	15.2	27.7	22.2	22.9	20.3	16.5	14.9	18.4	14.3
59.6	44.5	18.3	15.6	21.4	17.3	24.1	24.3	15.3	16.2	11.1	10.4
39.0	21.7	18.7	16.5	31.1	29.4	39.6	27.2	30.5	28.4	23.4	20.7
45.1	25.2	12.2	8.9	24.0	20.8	29.0	18.6	32.0	26.6	43.8	34.2
21.5	16.4										

Evaluation of uncorrected data for CM1

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	0.793 *	random term	3.48	µg/m³
uncertainty of b	0.024	bias at LV	-10.25	µg/m³
intercept a	0.09	combined uncertainty	10.83	µg/m³
uncertainty of a	0.81	relative uncertainty	21.7%	fail
number of datapairs	157	RM between-sampler uncertainty	1.50	µg/m³

* Significant at 95% confidence level.

Evaluation of data for CM1 after correction for slope

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	1.018	random term	4.74	µg/m³
uncertainty of b	0.030	bias at LV	0.44	µg/m³
intercept a	-0.44	combined uncertainty	4.76	µg/m³
uncertainty of a	1.03	relative uncertainty	9.5%	pass
number of datapairs	157	RM between-sampler uncertainty	1.50	µg/m³

Comparison of CM2 and RM – location E

Full dataset (µg.m^{-3}, uncorrected)

RM	*CM2*	*RM*	*CM2*	*RM*	*CM2*	*RM*	*CM2*	*RM*	*CM2*	*RM*	*CM2*
10.0	9.8	44.3	36.8	22.1	21.9	35.3	33.6	23.7	23.9	48.1	42.9
21.8	17.0	37.0	35.7	12.1	12.2	20.6	17.7	34.1	23.5	47.6	33.2
31.2	22.4	98.4	79.9	18.4	18.0	24.7	23.8	20.6	19.6	18.8	17.0
35.0	26.9	78.5	58.1	17.7	19.3	24.0	22.4	34.5	26.9	33.2	19.1
66.9	51.7	67.8	55.2	20.1	21.4	29.0	27.6	17.9	20.8	19.7	16.4
72.1	52.7	28.3	25.3	19.1	18.8	20.0	18.9	49.2	41.4	31.3	25.5
35.0	26.2	13.6	11.8	23.4	24.0	24.2	21.5	57.2	54.8	27.1	27.7
38.7	29.5	34.6	28.5	29.7	26.2	19.2	19.7	45.8	44.3	32.9	28.3
29.6	23.4	28.3	24.0	34.9	32.6	37.8	36.5	44.0	39.2	38.9	26.3
10.7	10.2	30.8	27.9	20.1	16.0	23.7	22.9	14.5	13.7	31.1	21.3
10.8	9.5	36.7	28.7	27.0	25.2	30.9	29.5	21.5	21.8	21.6	18.2
30.7	27.4	52.0	45.0	28.5	25.6	46.8	44.2	21.2	20.9	75.0	60.4
43.3	32.7	43.0	37.3	25.1	22.1	54.9	53.7	18.9	16.1	35.4	28.8
13.8	14.0	44.5	41.8	49.1	36.5	62.4	59.1	26.7	21.5	21.5	16.1
26.2	20.6	28.7	27.3	24.2	17.9	47.7	47.9	42.6	33.3	96.4	81.2
16.0	15.6	38.1	34.2	29.7	26.6	19.8	18.5	58.2	43.9	40.3	31.4
63.8	53.6	56.4	55.8	26.8	24.3	20.9	21.2	33.0	25.2	7.2	6.9
45.0	40.8	22.8	21.0	24.5	22.2	27.4	24.2	39.0	30.5	13.4	13.8
38.7	32.9	16.6	13.5	51.4	41.0	44.9	39.6	21.1	15.3	54.4	45.5
56.7	50.9	25.6	24.9	41.7	35.3	27.8	22.3	82.9	69.8	37.0	28.3
61.1	53.4	13.9	12.5	18.3	17.1	24.6	19.2	22.8	15.8	15.8	16.5
58.0	51.9	16.4	14.6	18.9	16.6	20.7	17.0	17.7	11.9	23.4	18.6
85.4	75.0	22.0	17.5	20.5	12.4	24.3	17.2	12.2	12.6	16.3	12.6
18.3	16.8	37.9	29.0	48.1	42.6	15.3	12.0	43.7	38.3	9.3	8.9
37.0	30.0	47.7	36.4	17.0	14.1	15.5	14.3	35.5	24.8	18.1	17.1
79.1	66.2	21.8	18.3	21.1	18.9	27.5	26.8	29.8	23.4	51.0	41.6
52.8	46.0	15.0	13.9	23.6	19.4						

Evaluation of uncorrected data for CM2

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	0.829 *	random term	2.70	µg/m³
uncertainty of b	0.014	bias at LV	-7.67	µg/m³
intercept a	0.88	combined uncertainty	8.13	µg/m³
uncertainty of a	0.52	relative uncertainty	16.3%	fail
number of datapairs	159	RM between-sampler uncertainty	1.50	µg/m³

* Significant at 95% confidence level.

Evaluation of data for CM2 after correction for slope

RESULTS OF REGRESSION		*EQUIVALENCE TEST RESULTS*		
slope b	1.004	random term	3.49	µg/m³
uncertainty of b	0.017	bias at LV	1.13	µg/m³
intercept a	0.93	combined uncertainty	3.67	µg/m³
uncertainty of a	0.63	relative uncertainty	7.3%	pass
number of datapairs	159	RM between-sampler uncertainty	1.50	µg/m³

European Commission

EUR 23216 EN– Joint Research Centre – Institute for Environment and Sustainability
Title: Demonstration of Equivalence of Ambient Air Monitoring Methods: PROCEEDINGS of the WORKSHOP on Demonstration of Equivalence. 2-4 May 2007 Ispra (VA) Italy. GUIDELINE on Demonstration of Equivalence. Novembre 2005.

Edited by: P. Pérez Ballesta
Luxembourg: Office for Official Publications of the European Communities
2008 – 179 pp. – 21.59 x 28.94 cm
EUR – Scientific and Technical Research series – ISSN 1018-5593
ISBN 978-92-79-08205-4
DOI 10.2788/63342

Abstract

This publication collates the proceedings of the Workshop on Demonstration of Equivalence between Ambient Air Monitoring Methods, which was held at the JRC in Ispra (VA) from 2 - 4 May 2007 and the major discussions vis-à-vis the last version of the “Guideline on demonstration of Equivalence”.

The extensive use by national reference laboratories of this guideline strengthens the principles on which the current guideline is based. The original presentations of this workshop can be downloaded from the web site of the JRC at http://ies.jrc.cec.eu.int/626.html.

How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

JRC
EUROPEAN COMMISSION

LB-NA-23216-EN-C

ISBN 978-92-79-08205-4